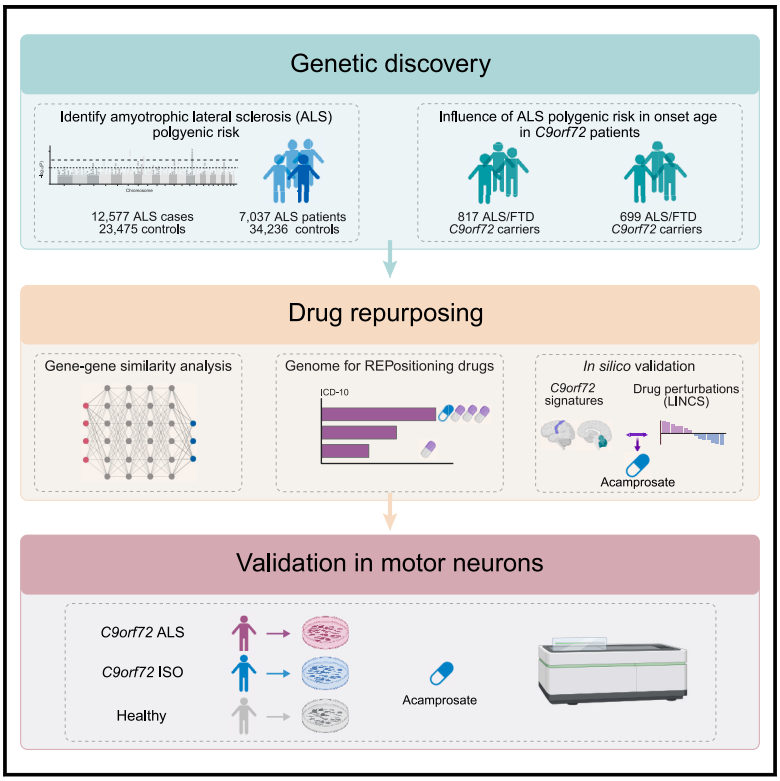
Cell Genomics

Article

Mechanism-free repurposing of drugs for *C9orf72*-related ALS/FTD using large-scale genomic data

Graphical abstract



Highlights

- Repeat expansions in *C9orf72* are the most common genetic cause of ALS and FTD
- The genetic risk of general ALS modifies the age at onset in *C9orf72* cases
- We performed a drug screen based on the genetic variants influencing age at onset
- We identified acamprosate, a GABA analog, as a potential treatment for *C9orf72*

Authors

Sara Saez-Atienzar, Cleide dos Santos Souza, Ruth Chia, ..., Laura Ferraiuolo, Isabella Fogh, Bryan J. Traynor

Correspondence

sara.saezatienzar@osumc.edu (S.S.-A.), traynorb@mail.nih.gov (B.J.T.)

In brief

Saez-Atienzar et al. identify genetic variants influencing the age at onset among patients carrying *C9orf72* repeat expansions. A drug screen based on these variants revealed acamprosate, a GABA analog, as a potentially repurposable treatment for *C9orf72*-related disease. The work underscores the potential of leveraging large-scale genomic data for drug repurposing.

Saez-Atienzar et al., 2024, Cell Genomics *4*, 100679
November 13, 2024 Published by Elsevier Inc.
https://doi.org/10.1016/j.xgen.2024.100679

CellPress

Cell Genomics

CellPress
OPEN ACCESS

Article

Mechanism-free repurposing of drugs for *C9orf72*-related ALS/FTD using large-scale genomic data

Sara Saez-Atienzar,[1,2,80,*] Cleide dos Santos Souza,[3,80] Ruth Chia,[1,80] Selina N. Beal,[3] Ileana Lorenzini,[4] Ruili Huang,[5] Jennifer Levy,[4] Camelia Burciu,[4] Jinhui Ding,[6] J. Raphael Gibbs,[6] Ashley Jones,[7] Ramita Dewan,[1] Viviana Pensato,[8] Silvia Peverelli,[9] Lucia Corrado,[10] Joke J.F.A. van Vugt,[11] Wouter van Rheenen,[11] Ceren Tunca,[12] Elif Bayraktar,[12] Menghang Xia,[5] The International ALS Genomics Consortium,[75] ITALSGEN Consortium,[76] SLAGEN Consortium,[77] Project MinE ALS Sequencing Consortium,[78] Alfredo Iacoangeli,[7,13,14] Aleksey Shatunov,[7] Cinzia Tiloca,[9] Nicola Ticozzi,[9,15] Federico Verde,[9,15] Letizia Mazzini,[16] Kevin Kenna,[11] Ahmad Al Khleifat,[7] Sarah Opie-Martin,[7] Flavia Raggi,[17]

(Author list continued on next page)

[1]Neuromuscular Diseases Research Section, National Institute on Aging, National Institutes of Health (NIH), Bethesda, MD 20892, USA
[2]Department of Neurology, Ohio State University, Columbus, OH 43210, USA
[3]Sheffield Institute for Translational Neuroscience, University of Sheffield, Sheffield S10 2HQ, UK
[4]Department of Translational Neuroscience, Barrow Neurological Institute, Phoenix, AZ 85013, USA
[5]Division of Pre-clinical Innovation, National Center for Advancing Translational Sciences (NCATS), NIH, Rockville, MD 20850, USA
[6]Computational Biology Group, Laboratory of Neurogenetics, National Institute on Aging, Bethesda, MD 20892, USA
[7]Maurice Wohl Clinical Neuroscience Institute, Department of Basic and Clinical Neuroscience, Institute of Psychiatry, Psychology, and Neuroscience, King's College London, London, UK
[8]Unit of Medical Genetics and Neurogenetics, Fondazione IRCCS Istituto Neurologico Carlo Besta, Milan, Italy
[9]Department of Neurology and Laboratory of Neuroscience, Istituto Auxologico Italiano IRCCS, Milan, Italy
[10]Department of Health Sciences, University of Eastern Piedmont, Novara, Italy
[11]Department of Neurology, UMC Utrecht Brain Center, University Medical Center Utrecht, Utrecht University, Utrecht, the Netherlands
[12]Neurodegeneration Research Laboratory (NDAL), Research Center for Translational Medicine (KUTTAM), Koç University School of Medicine, Istanbul, Turkey
[13]Department of Biostatistics and Health Informatics, Institute of Psychiatry, Psychology, and Neuroscience, King's College London, London, UK

(Affiliations continued on next page)

SUMMARY

Repeat expansions in the *C9orf72* gene are the most common genetic cause of (ALS) and frontotemporal dementia (FTD). Like other genetic forms of neurodegeneration, pinpointing the precise mechanism(s) by which this mutation leads to neuronal death remains elusive, and this lack of knowledge hampers the development of therapy for *C9orf72*-related disease. We used an agnostic approach based on genomic data (n = 41,273 ALS and healthy samples, and n = 1,516 *C9orf72* carriers) to overcome these bottlenecks. Our drug-repurposing screen, based on gene- and expression-pattern matching and information about the genetic variants influencing onset age among *C9orf72* carriers, identified acamprosate, a γ-aminobutyric acid analog, as a potentially repurposable treatment for patients carrying *C9orf72* repeat expansions. We validated its neuroprotective effect in cell models and showed comparable efficacy to riluzole, the current standard of care. Our work highlights the potential value of genomics in repurposing drugs in situations where the underlying pathomechanisms are inherently complex.

INTRODUCTION

A repeat expansion within the *C9orf72* gene is a common cause of amyotrophic lateral sclerosis (ALS) and frontotemporal dementia (FTD), two neurological disorders that result in the deaths of ~17,000 Americans and Europeans annually.[1–4] This genetic disease accounts for 1 in 10 ALS and FTD cases of European descent, and many carriers in the general population have a near-complete chance of manifesting symptoms during their lifetime.[5,6] The exact processes by which this repeat expansion leads to neuronal death are not fully understood, although several mechanisms, such as dipeptide production, RNA

Massimiliano Filosto,[18,19] Stefano Cotti Piccinelli,[18,19] Alessandro Padovani,[19] Stella Gagliardi,[20] Maurizio Inghilleri,[21,22] Alessandra Ferlini,[23] Rosario Vasta,[24] Andrea Calvo,[24,25] Cristina Moglia,[24,25] Antonio Canosa,[24,25,26] Umberto Manera,[24,25] Maurizio Grassano,[24] Jessica Mandrioli,[27,28] Gabriele Mora,[24] Christian Lunetta,[29,30] Raffaella Tanel,[31] Francesca Trojsi,[32] Patrizio Cardinali,[33] Salvatore Gallone,[24] Maura Brunetti,[24] Daniela Galimberti,[34,35] Maria Serpente,[34] Chiara Fenoglio,[34,35] Elio Scarpini,[34] Giacomo P. Comi,[15,36] Stefania Corti,[15,36] Roberto Del Bo,[15,36] Mauro Ceroni,[20,37] Giuseppe Lauria Pinter,[38,39] Franco Taroni,[8] Eleonora Dalla Bella,[38] Enrica Bersano,[38,40]

(Author list continued on next page)

[14]National Institute for Health Research Biomedical Research Centre and Dementia Unit, South London and Maudsley NHS Foundation Trust and King's College London, London, UK
[15]Department of Pathophysiology and Transplantation, "Dino Ferrari" Center, Università degli Studi di Milano, Milan, Italy
[16]Amyotrophic Lateral Sclerosis Center, Department of Neurology "Maggiore della Carità" University Hospital, Novara, Italy
[17]Department of Neurosciences, University of Padova, Padova, Italy
[18]NeMO-Brescia Clinical Center for Neuromuscular Diseases, University of Brescia, Brescia, Italy
[19]Department of Clinical and Experimental Sciences, University of Brescia, Brescia, Italy
[20]Genomic and Post-Genomic Center, IRCCS Mondino Foundation, Pavia, Italy
[21]Department of Human Neurosciences, Rare Neuromuscular Diseases Centre, Sapienza University, 00185 Rome, Italy
[22]IRCCS Neuromed, Pozzilli, Italy
[23]Unit of Medical Genetics, Department of Medical Science, University of Ferrara, Ferrara, Italy
[24]"Rita Levi Montalcini" Department of Neuroscience, Amyotrophic Lateral Sclerosis Center, University of Turin, Turin, Italy
[25]Azienda Ospedaliero Universitaria Città della Salute e della Scienza, Turin, Italy
[26]Institute of Cognitive Sciences and Technologies, C.N.R., Rome, Italy
[27]Department of Biomedical, Metabolic, and Neural Sciences, University of Modena and Reggio Emilia, Modena, Italy
[28]Department of Neurosciences, Azienda Ospedaliero Universitaria di Modena, Modena, Italy
[29]Department of Neurorehabilitation, Istituti Clinici Scientifici Maugeri IRCCS, Institute of Milan, Milan, Italy
[30]NEMO Clinical Center Milano, Fondazione Serena Onlus, Milan, Italy
[31]Operative Unit of Neurology, S. Chiara Hospital, Trento, Italy
[32]Department of Advanced Medical and Surgical Sciences, University of Campania "Luigi Vanvitelli," Naples, Italy
[33]Neurology Unit, AST Fermo, Marche, Italy
[34]Neurodegenerative Diseases Unit, Fondazione IRCCS Ca' Granda, Ospedale Maggiore Policlinico, Milan, Italy
[35]Department of Biomedical, Surgical, and Dental Sciences, University of Milan, Milan, Italy
[36]Neurology Unit, Fondazione IRCCS Ca' Granda, Ospedale Maggiore Policlinico, Milan, Italy
[37]Department of Brain and Behavioural Sciences, University of Pavia, Pavia, Italy
[38]3rd Neurology Unit, Motor Neuron Diseases Center, Fondazione IRCCS Istituto Neurologico "Carlo Besta," Milan, Italy
[39]Department of Medical Biotechnology and Translational Medicine, University of Milan, Milan, Italy
[40]"L. Sacco" Department of Biomedical and Clinical Sciences, Università degli Studi di Milano, Milan, Italy
[41]Social Genetic & Developmental Psychiatry Centre, Institute of Psychiatry, Psychology, and Neuroscience (IoPPN), King's College London, London, UK
[42]NIHR BioResource Centre Maudsley, NIHR Maudsley Biomedical Research Centre (BRC) at South London and Maudsley NHS Foundation Trust (SLaM), London, UK
[43]School of Medicine, Dentistry, and Biomedical Sciences, Faculty of Medicine Health and Life Sciences, Queen's University, Belfast, UK
[44]Sheffield Institute for Translational Neuroscience (SITraN), University of Sheffield, and the NIHR Sheffield Biomedical Research Centre, Sheffield, UK
[45]Department of Biostatistics and Health Informatics, Institute of Psychiatry, Psychology, and Neuroscience (IoPPN), King's College London, London SE5 8AF, UK
[46]NIHR Biomedical Research Centre at South London and Maudsley NHS Foundation Trust and King's College London, London, UK
[47]Health Data Research UK London, University College London, London, UK
[48]Institute of Health Informatics, University College London, London, UK
[49]NIHR Biomedical Research Centre at University College London Hospitals NHS Foundation Trust, London, UK
[50]Department of Anatomy, Physiology, & Genetics, Uniformed Services University of the Health Sciences, Bethesda, MD 20814, USA
[51]The American Genome Center, Uniformed Services University of the Health Sciences, Bethesda, MD 20814, USA
[52]Neurodegenerative Diseases Research Section, National Institute of Neurological Disorders and Stroke (NINDS), NIH, Bethesda, MD 20892, USA
[53]Department of Neurology, Johns Hopkins University Medical Center, Baltimore, MD 21287, USA
[54]Department of Clinical Neuroscience, King's College Hospital, London SE5 9RS, UK
[55]Complex Trait Genomics Laboratory, Smurfit Institute of Genetics, Trinity College Dublin, Dublin, Ireland
[56]Academic Unit of Neurology, Trinity Biomedical Sciences Institute, Trinity College Dublin, Dublin, Ireland
[57]Department of Neurology, Beaumont Hospital, Dublin, Ireland
[58]Euan MacDonald Centre for Motor Neurone Disease Research, Edinburgh, UK
[59]UK Dementia Research Institute, University of Edinburgh, Edinburgh, UK
[60]Centre for Neuroregeneration and Medical Research Council Centre for Regenerative Medicine, University of Edinburgh, Edinburgh, UK

(Affiliations continued on next page)

Charles J. Curtis,[41,42] Sang Hyuck Lee,[41,42] Raymond Chung,[41,42] Hamel Patel,[13,42] Karen E. Morrison,[43] Johnathan Cooper-Knock,[44] Pamela J. Shaw,[44] Gerome Breen,[41,42] Richard J.B. Dobson,[45,46,47,48,49] Clifton L. Dalgard,[50,51] The American Genome Center[79], Sonja W. Scholz,[52,53] Ammar Al-Chalabi,[7,54] Leonard H. van den Berg,[11] Russell McLaughlin,[55] Orla Hardiman,[56,57] Cristina Cereda,[20] Gianni Sorarù,[17] Sandra D'Alfonso,[10] Siddharthan Chandran,[58,59] Suvankar Pal,[58,60] Antonia Ratti,[9,61] Cinzia Gellera,[8] Kory Johnson,[62] Tara Doucet-O'Hare,[63] Nicholas Pasternack,[64] Tongguang Wang,[64] Avindra Nath,[64] Gabriele Siciliano,[65] Vincenzo Silani,[9,15] Ayşe Nazlı Başak,[12] Jan H. Veldink,[11] William Camu,[66] Jonathan D. Glass,[67] John E. Landers,[68] Adriano Chiò,[24,25,26] Rita Sattler,[4] Christopher E. Shaw,[69,70,80] Laura Ferraiuolo,[71,80] Isabella Fogh,[69,80] and Bryan J. Traynor[1,53,72,73,74,80,81,*]

[61]Department of Medical Biotechnology and Translational Medicine, Università degli Studi di Milano, Milan, Italy
[62]Bioinformatics Section, Information Technology Program (ITP), Division of Intramural Research (DIR), National Institute of Neurological Disorders & Stroke, NIH, Bethesda, MD 20892, USA
[63]Neuro-oncology Branch, Center for Cancer Research, National Cancer Institute (NCI), NIH, Bethesda, MD 20892, USA
[64]Translational Neuroscience Center, National Institute of Neurological Disorders and Stroke (NINDS), NIH, Bethesda, MD 20892, USA
[65]Department of Clinical and Experimental Medicine, University of Pisa, Pisa, Italy
[66]ALS Center, CHU Gui de Chauliac, University of Montpellier, Montpellier, France
[67]Department of Neurology, Emory University School of Medicine, Atlanta, GA 30322, USA
[68]Department of Neurology, University of Massachusetts Medical School, Worcester, MA, USA
[69]United Kingdom Dementia Research Institute, Maurice Wohl Clinical Neuroscience Institute, Department of Basic and Clinical Neuroscience, Institute of Psychiatry, Psychology, and Neuroscience, King's College London, London, UK
[70]Centre for Brain Research, University of Auckland, Auckland, New Zealand
[71]Sheffield Institute for Translational Neuroscience, University of Sheffield, Sheffield S10 2HQ, UK
[72]Reta Lila Weston Institute, UCL Queen Square Institute of Neurology, University College London, London WC1N 1PJ, UK
[73]National Institute of Neurological Disorders and Stroke (NINDS), NIH, Bethesda, MD 20892, USA
[74]RNA Therapeutics Laboratory, National Center for Advancing Translational Sciences (NCATS), NIH, Rockville, MD 20850, USA
[75]International ALS Genomics Consortium members are listed fully in the supplemental information.
[76]ITALSGEN Consortium members are listed fully in the supplemental information.
[77]SLAGEN Consortium members are listed fully in the supplemental information.
[78]Project MinE ALS Sequencing Consortium members are listed fully in the supplemental information.
[79]The American Genome Center members are listed fully in the supplemental information.
[80]These authors contributed equally
[81]Lead contact
*Correspondence: sara.saezatienzar@osumc.edu (S.S.-A.), traynorb@mail.nih.gov (B.J.T.)
https://doi.org/10.1016/j.xgen.2024.100679

toxicity, disruption of nucleocytoplasmic transport, and haploinsufficiency, have been suggested.[7] This lack of molecular knowledge is a common theme across neurodegenerative disorders, where there is often a complex interplay among multiple pathways and cellular processes.[8–11]

The molecular complexity underlying neurodegenerative diseases also hampers drug discovery; ameliorating a single aspect of a cellular network may not be beneficial as it does not address the other pathological processes co-occurring within the cell.[12] Traditional linear drug-discovery efforts, where millions of compounds are tested against a single target, are likely to fail in the face of such multidimensional conditions, and this disconnect alone may account for the high failure rate observed among clinical trials in neurodegenerative diseases.[13] Genomic and transcriptomic data offer a potential solution, as these data types inherently capture the multifaceted nature of neurological diseases. We can exploit this information to match drugs that restore entire networks and systems, even when the target pathways or the mechanism of action of the drug are not fully understood.[14,15] Compounds supported by genetic evidence are also more likely to succeed in clinical trials and to gain drug approval.[16,17]

In this context, we have used a massive genomic dataset to identify the genetic variants influencing age at onset among patients carrying the *C9orf72* repeat expansion. Building on this genomic information, we have leveraged gene- and expression-pattern matching and pathway modeling to nominate potentially repurposable drugs for *C9orf72*-related ALS/FTD. We focused on onset age because of the wide range observed in this common genetic form of neurodegeneration, spanning from the fourth to the tenth decade of life, representing a natural experiment within the ALS population.[5,18] Genetic factors are known to play a role in this variable age-at-onset presentation,[19] and research in other neurological diseases shows that the phenotypic manifestations of high-risk pathogenic variants are influenced by minor effect variants elsewhere in the genome.[20,21]

Our pipeline nominated acamprosate, an oral medication used to manage alcohol use disorder, as a potential repurposable drug for slowing progression among symptomatic individuals and delaying disease onset among *C9orf72* carriers. *In vitro* data demonstrated a neuroprotective effect of acamprosate in motor neurons derived from these patients. Crucially, our discovery approach evolved from the notion that genomics can nominate specific medications in an agnostic, data-driven manner, even when precise knowledge about disease mechanisms or drug pharmacodynamics is lacking.

RESULTS

Figure 1 shows a graphical representation of our genetic workflow and drug-repurposing pipeline.

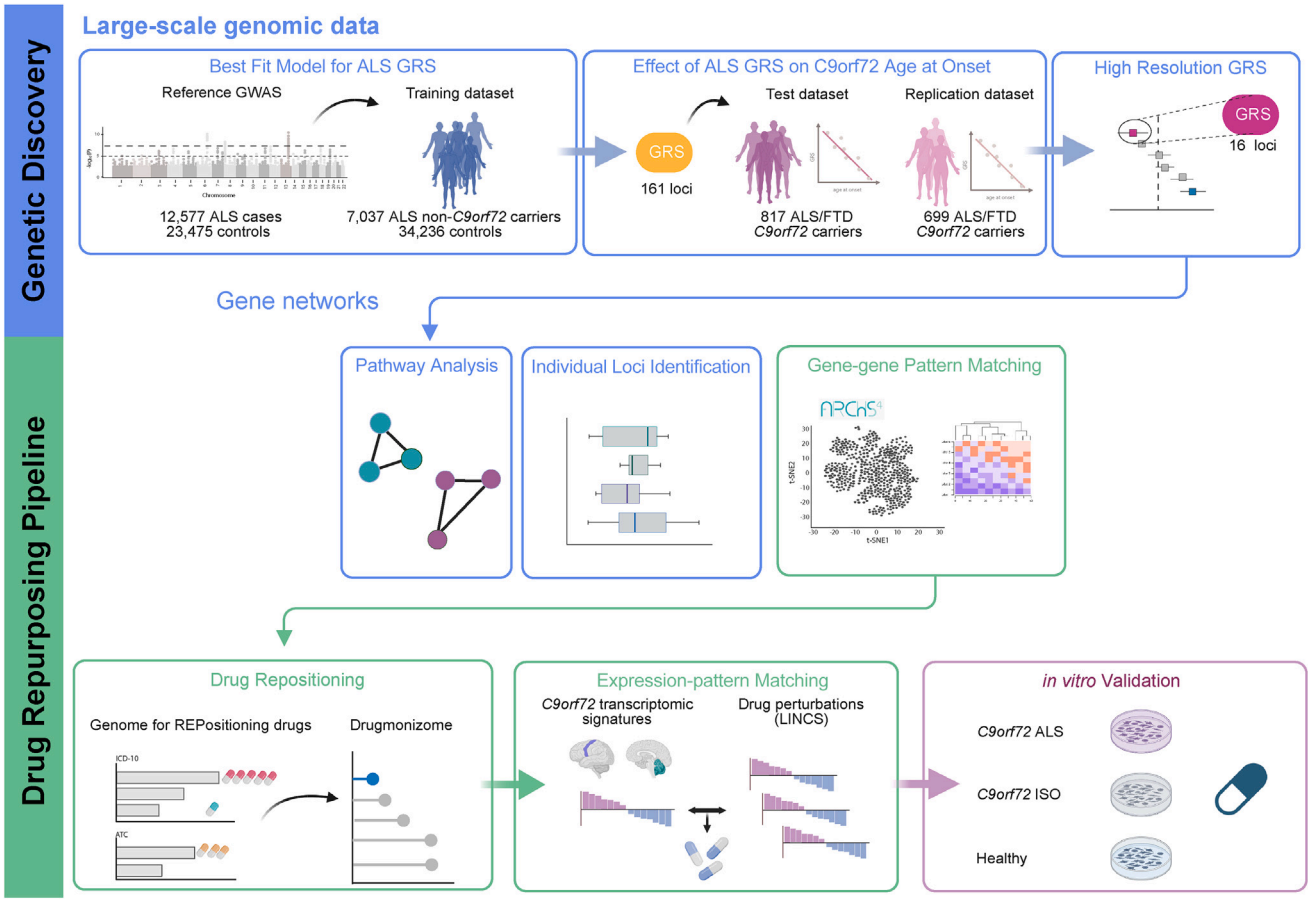


Figure 1. Schematic illustration of the analytical workflow

The genetic risk score for sporadic ALS was generated using large cohorts as the reference and training sets. This general ALS genetic risk score was then calculated for a sizable cohort of *C9orf72* carriers. Follow-up analyses included pathway analysis and the identification of individual loci with a major contribution to the age at onset. Using the information obtained from these genetic analyses, we performed drug repurposing based on gene-gene-pattern matching and expression-pattern matching to identify drugs that may delay symptom onset among *C9orf72* carriers. *In vitro* drug validation confirmed the neuroprotective effect of the drug nominated by this approach.

Sporadic ALS risk variants modify the onset age among *C9orf72*-related ALS/FTD

We built a polygenic risk profile for general ALS using 7,037 cases not carrying the *C9orf72* repeat expansion and 34,236 healthy controls. The best-fit model included 161 SNPs (odds ratio for the model = 1.125, 95% confidence interval [CI] = 1.093–1.156, $p = 1.5 \times 10^{-16}$). Figure 2 and Table S1 describe the genetic variants and mapped genes that make up the polygenic risk score for general ALS.

Next, using the 161 SNPs, we investigated the association of the general ALS genetic risk score with age at onset in a cohort of 1,516 ALS/FTD *C9orf72* expansion carriers. We found the risk of general ALS (i.e., independent of the *C9orf72* repeat expansion) was significantly associated with onset age among the *C9orf72* expansion carriers (meta-analysis $p = 1.5 \times 10^{-3}$, $\beta = -0.781$, 95% CI = −1.263 to −0.299; Figures 3A–3C; Table S2). We observed similar results when analyzing ALS *C9orf72* and FTD *C9orf72* carrier patients alone (Figure S1). In contrast, the general ALS genetic risk score did not influence the age at onset among patients without *C9orf72* repeat expansions ($p = 0.437$, $\beta = 0.115$, 95% CI = −0.175 to 0.404; Figures 3C and S2; Table S2).

A subset of 16 SNPs drives the early-onset age among *C9orf72*-related ALS/FTD

Having established that the general ALS risk influenced the age at onset among *C9orf72* patients, we next determined which of the 161 SNPs in the model was driving the effect.[21] To do this, we performed a leave-one-out analysis to evaluate the contribution of each SNP to the onset age. This enabled us to rank the 161 SNPs and group the variants into 10 deciles (Table S3). We recalculated the genetic risk scores using the SNPs from each decile and used a linear regression model to test these refined genetic risk scores for association with age at onset.[22–27] Figure S3 shows a schematic representation of this approach.

Decile 10 contained the 16 SNPs with the most significant association with earlier onset age (Figures 4A and 4B). A

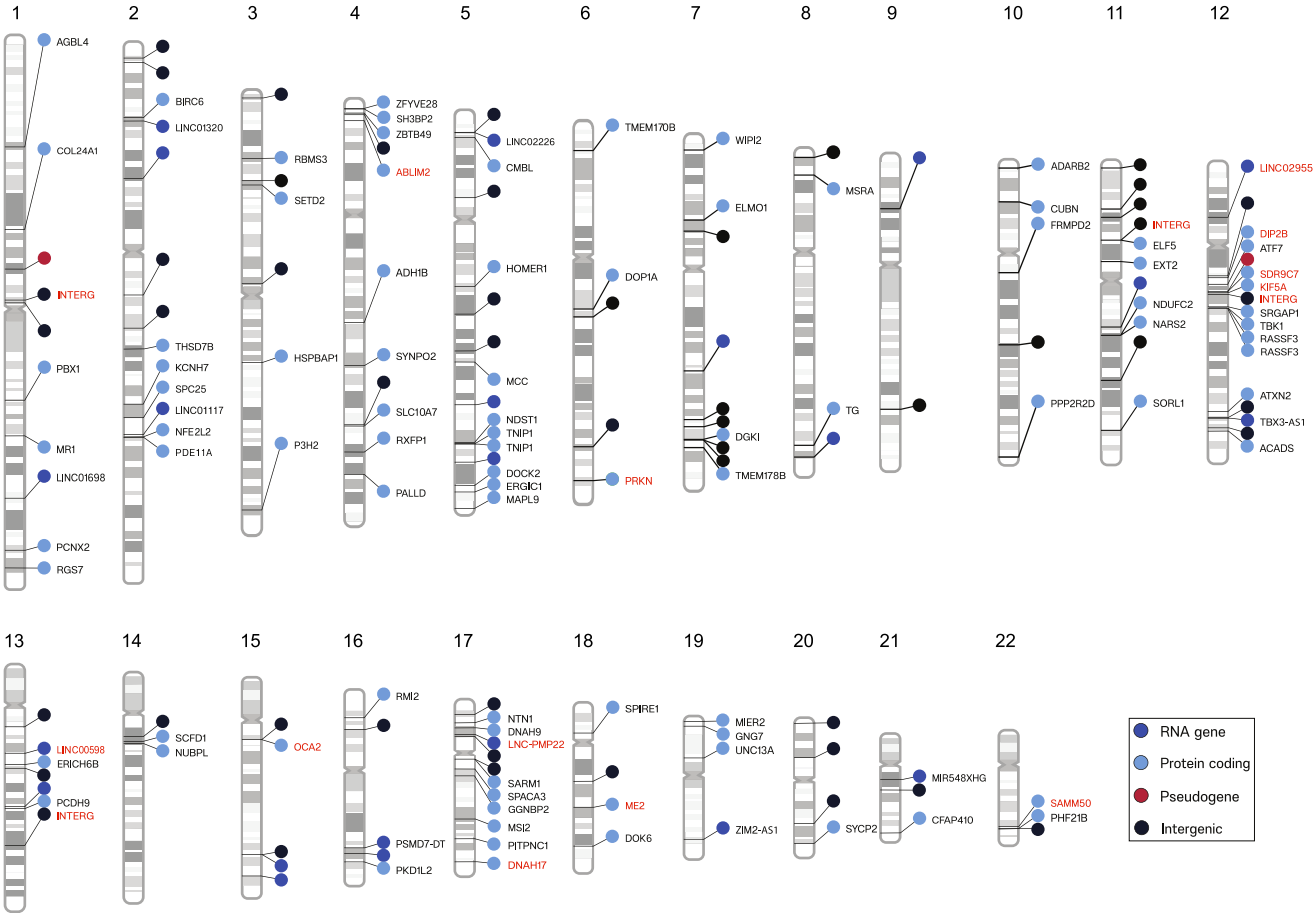


Figure 2. Genetic variants influencing symptom onset age among ALS/FTD patients carrying the *C9orf72* repeat expansion

The ideograms show the 161 ALS genetic risk loci making up the general ALS polygenic risk score. Labels with red text denote the 16 SNPs making up decile 10. The nearest genes to the variants are displayed. The colors of the circles correspond to the gene type: dark blue, RNA gene; light blue, protein-coding gene; red, pseudogene; black, intergenic. The numbers at the top indicate the chromosome. Interg, intergenic.
See also Table S1.

1-SD increase in the genetic risk score of the decile 10 SNPs corresponded to a decrease of 2.17 years (95% CI = 1.52–2.81) in the age at onset, and there was a 4-SD difference between *C9orf72* carriers at the extremes of the genetic risk score distribution. This difference implies that individuals at the highest end of the genetic risk score distribution developed the disease, on average, 8 years (8.68 years, 95% CI = 6.08–11.24; Figure 4B) earlier than their counterparts at the lowest end. In contrast, these loci did not alter the onset age among ALS patients not carrying the *C9orf72* repeat expansion (Figure S4).

Cytoskeletal and axonal transport pathways influence *C9orf72* age at onset

Enrichment analysis of the SNPs in decile 10 revealed that the cytoskeletal and axonal transport pathways influenced the age at onset among patients carrying the *C9orf72* repeat expansion (Figures 4C and S5; Table S4). These data suggested that these biological processes are essential in determining the onset of *C9orf72*-related disease.

The two-step design of our drug-repurposing pipeline

We used our genetic data to perform a drug-repositioning analysis to discover medications that may delay the age at disease onset among *C9orf72* patients. This approach harnesses our insights into the genetic factors influencing the variable age at onset in *C9orf72*-related disease and existing databases containing information on drug effects.[28] A critical advantage of our pipeline is that it identifies drugs with broad, heterogeneous effects for use in complex human diseases like neurodegenerative conditions. Crucially, this approach is not based on a single change within the cell but instead relies on gene-pattern and expression-pattern matching to select medications that correct the disrupted networks.

The pipeline comprises two complementary parts (see Figure 1 for a graphical representation of these steps). In the first step, we performed gene-drug-pattern matching. To do this, we used the *g:SNPense* function of g:Profiler2[29] to identify the genes related to the SNPs in decile 10. These genes were then used as search terms (defined as seed genes in Table S5) in the Geneshot web server (accessed October

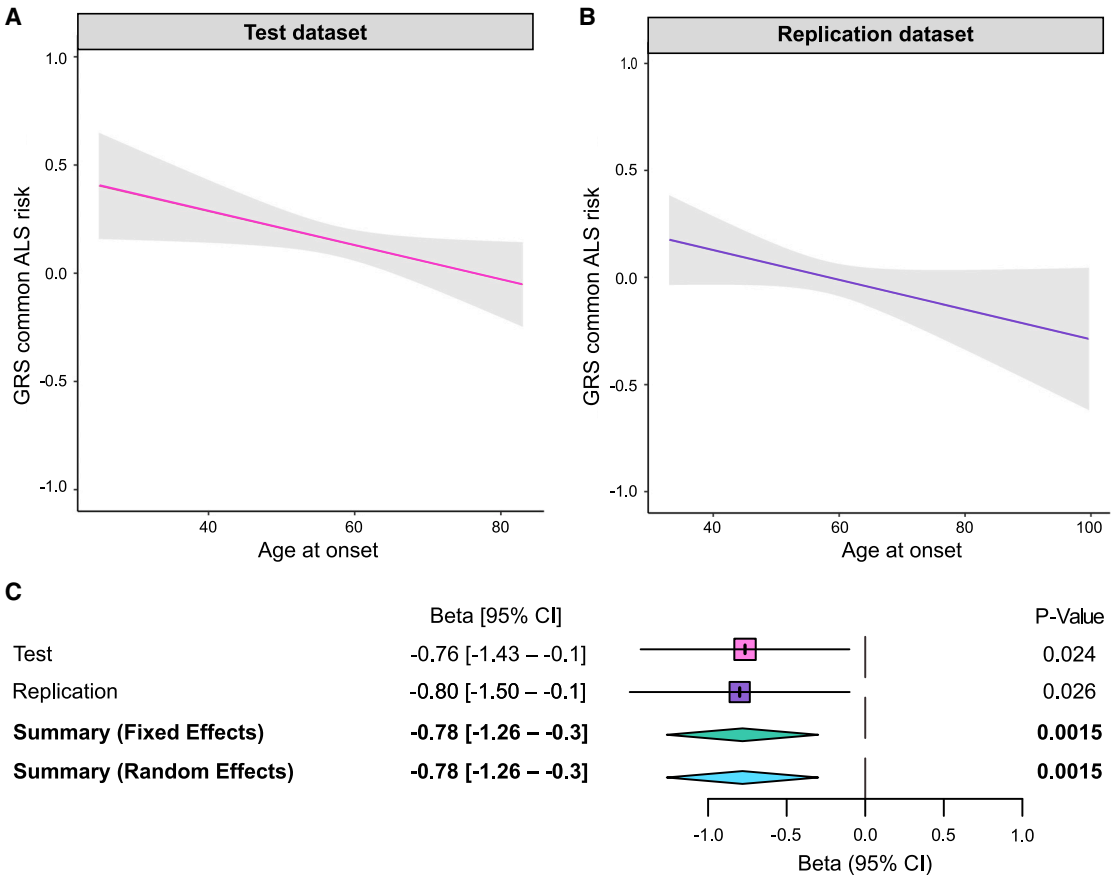


Figure 3. The ALS genetic risk score significantly influences onset age in *C9orf72* carriers

(A and B) The regression lines show the association of the ALS genetic risk scores and age at onset in (A) the test dataset (n = 817 ALS/FTD *C9orf72* carriers, p = 0.024, β = −0.765, 95% CI = −1.429 to −0.101) and (B) the replication dataset (n = 699, p = 0.026, β = −0.799, 95% CI = −1.499 to −0.099). The shadow areas represent the 90% confidence interval of the regression model.

(C) The forest plot shows the results of the meta-analysis of the test and replication datasets ($p = 1.5 \times 10^{-3}$, β = −0.781, 95% CI = −1.263 to −0.299).

See also Figures S1 and S2 and Table S2.

2021) to identify a broader list of functionally related genes.[30] This database identifies genes associated with the search term based on their co-occurrence in publications and gene-gene similarity from human RNA sequencing (RNA-seq) data (ARCHS4).[30] The resulting list of genes plus the seed genes (Table S5) was then used as input for the Genome for REPositioning drugs (GREP) analysis (version 1.0.0), which is a software package designed to identify drugs that target the gene set based on their enrichment in clinical indication categories.[31] The gene-drug software used Fisher's exact test to perform pharmacological enrichment and output the names of drugs associated with the gene set.

In the second step of our pipeline, we used gene expression-pattern matching to refine and narrow the candidate drug list. In essence, this approach identifies compounds that reverse the transcription patterns observed in brain tissue obtained from *C9orf72* patients (n = 44 cases and 76 healthy controls). This method is widely used in drug repositioning to assess how well drugs can counteract disease-related gene expression patterns by comparing their effects to a disease gene signature.[32] To do this, the drug perturbations were queried using the Library of Integrated Network-Based Cellular Signatures database,[33] containing the differential expression analysis of 12,328 genes from 8,140 compound treatments of 30 cell lines. A bidirectional weighted Kolmogorov-Smirnov enrichment statistic test of gene expression ranks in the disease and the expression values of the drug signatures was used to assign a connectivity map (CMap) score to each drug, reflecting the degree to which the drug "flips" the gene expression signature of the disease.

Step 1: Drug prioritization using gene-drug-pattern matching

The genomic-based GREP analysis nominated 52 medications approved for human use that could be repurposed to delay onset among *C9orf72* carriers (Figure 5A; Table S6). To investigate the targets of the nominated medications, we explored the significant biomedical terms associated with these drugs in the Drugmonizome database (accessed October 12, 2022).[34] These therapies were enriched for CNS targets and are typically prescribed for anxiety disorders and epilepsy; exploring their mechanisms

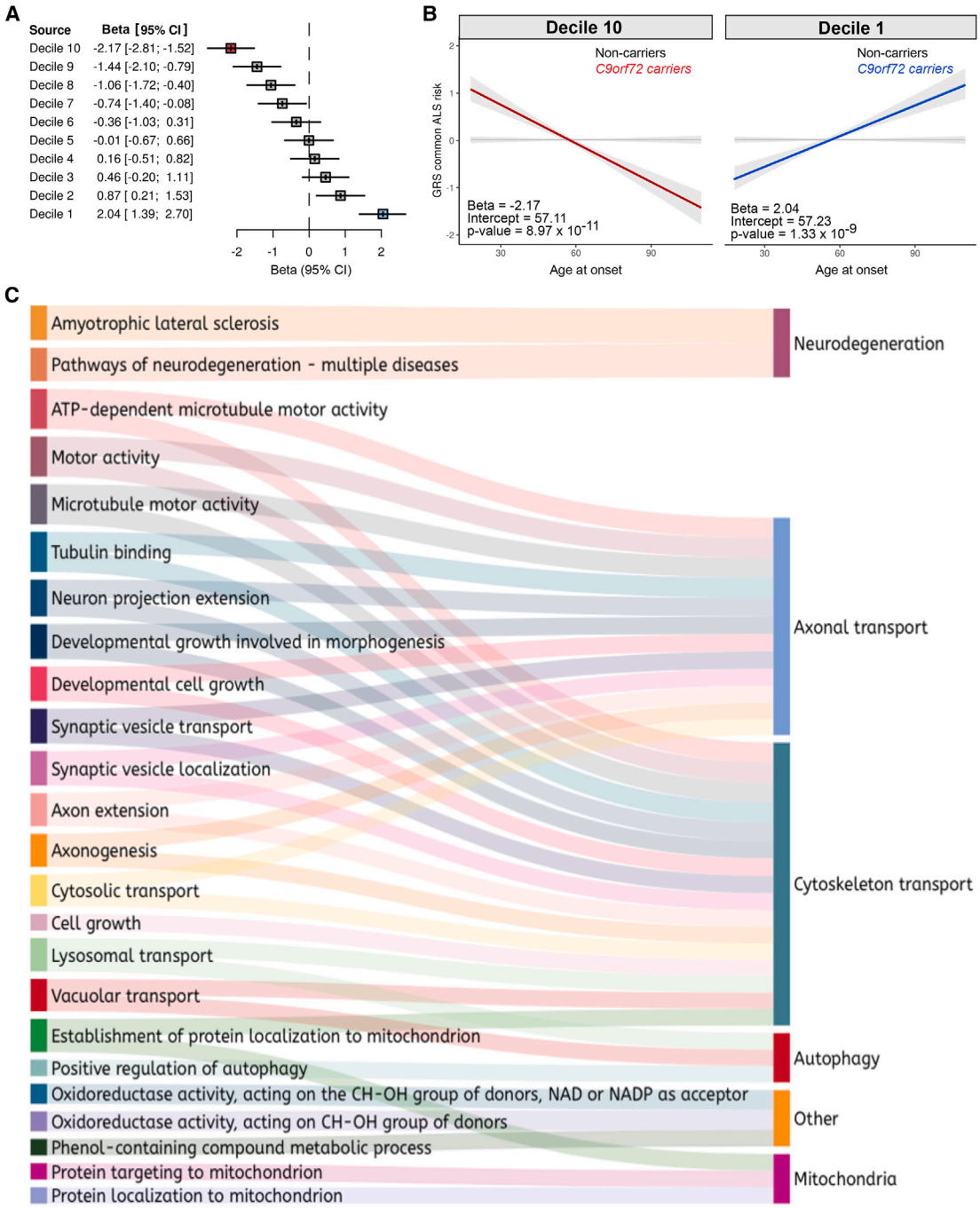


Figure 4. Contribution of individual SNPs to age at symptom onset among *C9orf72* patients

(A) The forest plot shows the effect size of each decile obtained by ranking the 161 individual SNPs based on their effect on age at onset in *C9orf72* patients.

(B) The regression lines show the association between the ALS genetic risk scores and age at onset in 817 ALS/FTD *C9orf72* carriers based on the 16 SNPs of decile 10 (n = 817 ALS/FTD *C9orf72* carriers, $p = 8.97 \times 10^{-11}$, $\beta = -2.17$, 95% CI = −2.81 to −1.52). The shadow area represents the 90% confidence interval of the regression model.

(C) Sankey diagram showing the functional enrichment of decile 10 genes, based on Gene Ontology terms. Only gene lists that contain between 5 and 500 genes were selected for the analysis. The significant threshold was an false discovery rate (FDR)-corrected $p < 0.05$.

See also Figures S3–S5 and Tables S3, S4, and S7.

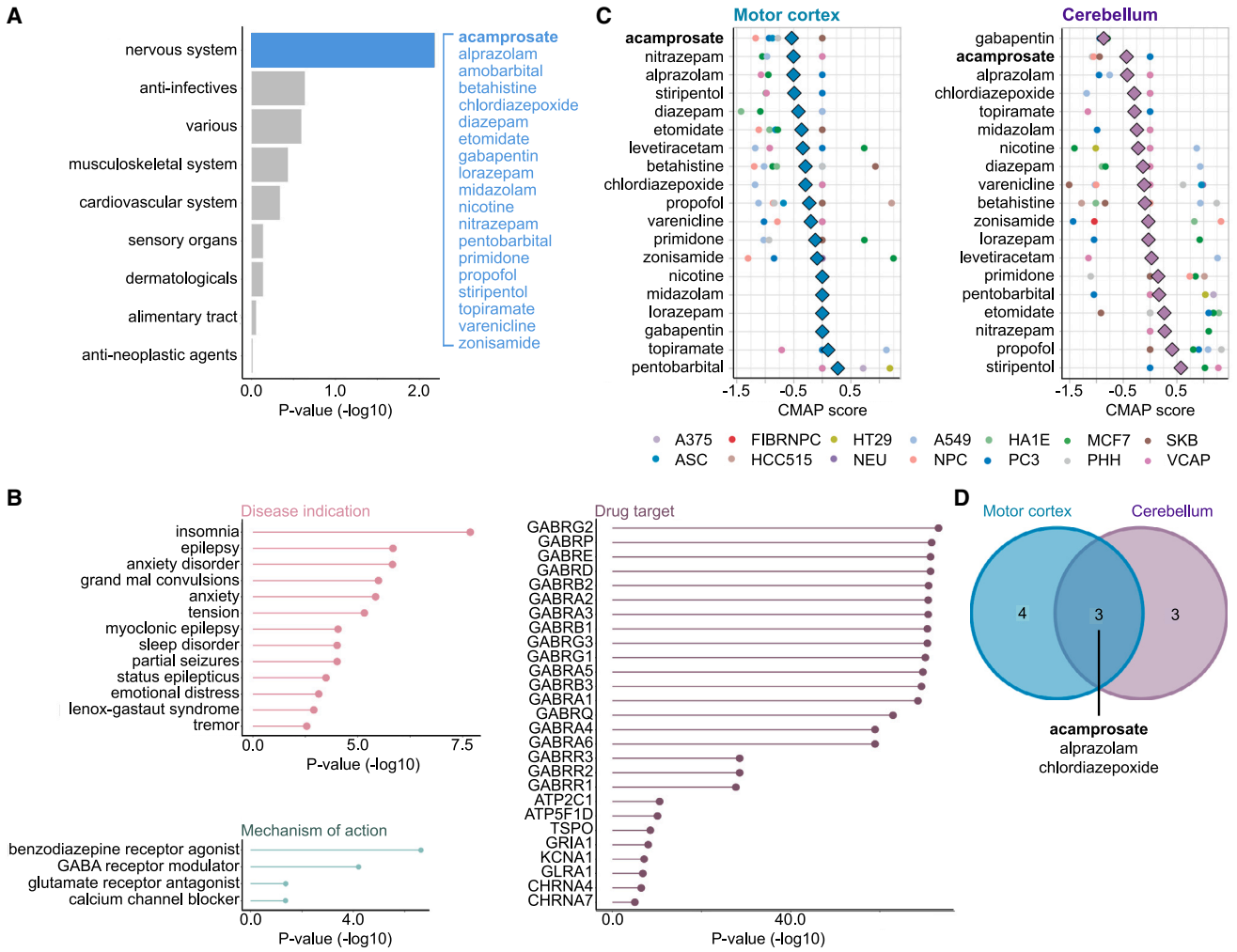


Figure 5. Repurposing drugs to delay onset among *C9orf72* carriers

The figure shows the results obtained from our drug-repositioning pipeline.

(A) Enriched terms derived from the GREP software package, which is based on the Anatomical Therapeutic Chemical classification. Blue indicates the significant "other nervous system drugs" category. Some of the drugs within this category are shown (see Table S6 for a complete list).

(B) Lollipop plots depict drug enrichment analysis for different categories, such as disease indication, gene target, and mechanism of action. Information was obtained from the Drugmonizome database, and the x axis depicts the enrichment corrected *p* value, which uses Bonferroni correction for the disease indication and the gene target plots and FDR for the mechanism of action.

(C) Drug perturbation data were obtained from the LINCS database. The graphs show the CMap scores for the selected drugs across cell types. CMap scores are determined using a bidirectional weighted Kolmogorov-Smirnov enrichment statistic test, which compares gene expression changes in the disease and drug signatures to quantify the extent to which the drug effectively reversed (flipped) the gene expression signature associated with the disease. Lower scores indicate a more substantial potential for therapeutic effectiveness. An average CMap (diamond shape) was calculated using the normalized connectivity score to evaluate the overall effect of each drug across the tested cell lines. Drugs with a reversal potential were selected if (1) they depicted a negative average CMap and (2) they showed a negative or neutral (measured as 0) CMap score for each cell line (circle shape). A375, ASC, FIBRNPC, HCC515, HT29, NEU, A549, NPC, HA1E, PC3, MCF7, PHH, SKB, and VCAP refer to the cell line types available in the LINCS database (see STAR Methods for details).

(D) The Venn diagram shows the drugs that fulfilled these criteria in the motor cortex and cerebellum. Of these, acamprosate was selected for additional *in vitro* validation.

See also Figures S5 and S6.

of action revealed the γ-aminobutyric acid (GABA) receptor as potentially relevant (Figure 5B).

Step 2: Drug prioritization using expression-pattern matching

The drug-disease expression-pattern matching Library of Integrated Network-Based Cellular Signatures (LINCS) analysis found that 3 of our 52 selected drugs demonstrated reversal of the *C9orf72* transcriptomic disease signature across multiple cell lines: acamprosate, chlordiazepoxide, and alprazolam (Figures 5C and 5D). Acamprosate was chosen from this list for the following reasons. First, acamprosate demonstrated neuroprotective properties in a SOD1^{G93A} rat spinal cord model of

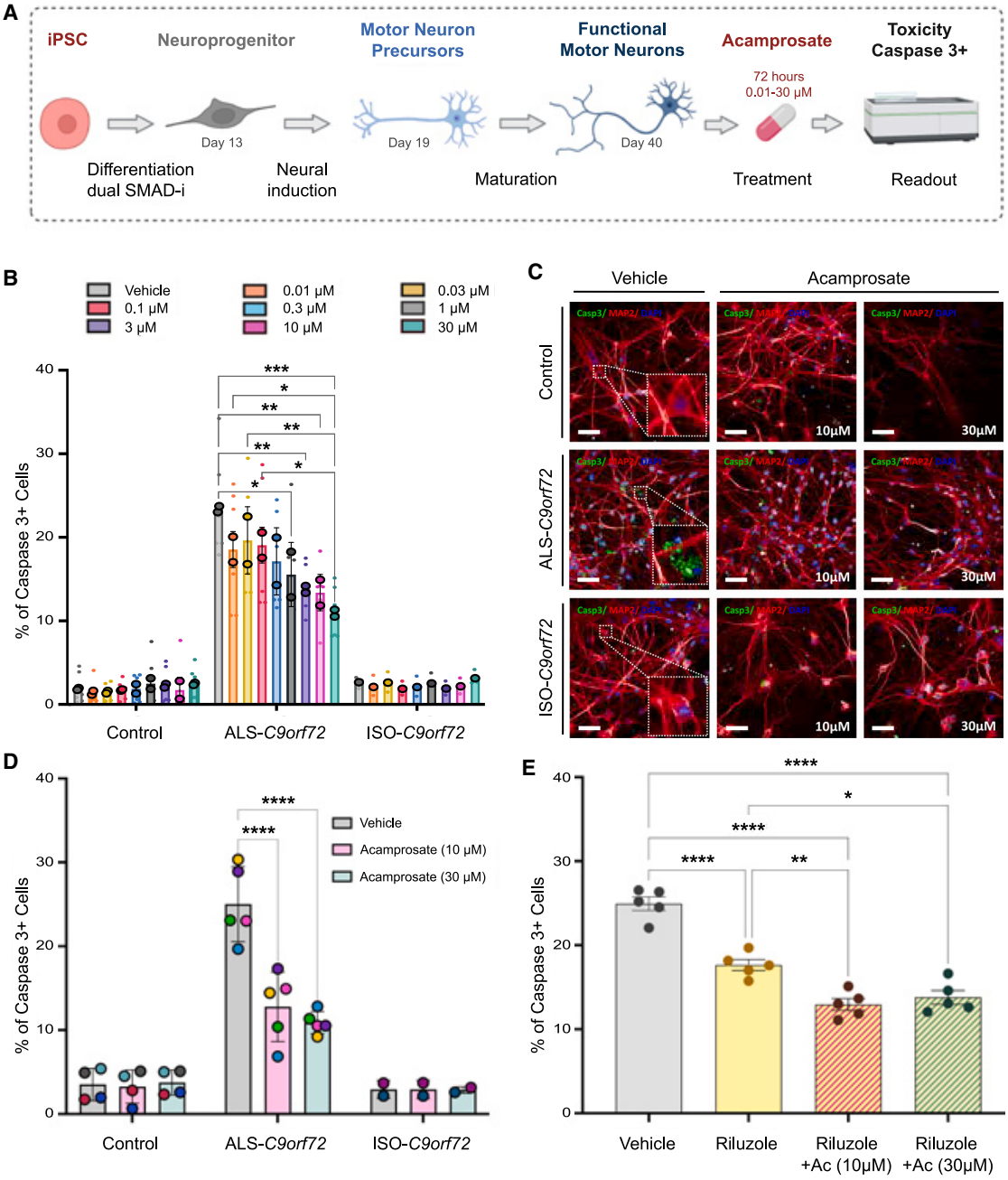


Figure 6. Acamprosate is neuroprotective in iPSC-derived motor neurons from *C9orf72* patients

(A) Schematic representation of the experiments to validate the effect of acamprosate.

(B) The bar graph depicts the percentage of cells showing cleaved caspase-3 (caspase-3^{+} cells) after acamprosate treatment. Minor dots represent biological replicates, averaging from three technical replicates each. In contrast, the bordered dots represent the mean effect in iPSC-derived motor neurons from two healthy donors, two *C9orf72* ALS patients, and one isogenic line. Data are mean ± SD. Comparisons within the control and the ALS-*C9orf72* groups were performed using a two-way ANOVA, with a Tukey's post hoc test ($n = 2$). Only $p < 0.05$ and comparisons with the vehicle group are displayed in the graph.

(C) Representative images of the motor neurons showing cleaved caspase-3 staining (green), MAP2 staining (red), and DAPI (blue). Scale bar, 50 μm. All molecular phenotypes were confirmed in a minimum of 3 technical replicates, and at least 25 fields were randomly selected and scanned per well of a 96-well plate in triplicate.

(D) Acamprosate effective doses (10 and 30 μM) were confirmed in additional iPSC-derived motor neurons, totaling cells from four healthy donors, five *C9orf72* ALS patients, and two isogenic lines. Each point represents the mean effect per cell line. Data are mean ± SD. Comparisons within the control, the ALS-*C9orf72*, and the ISO-*C9orf72* groups were performed using a two-way ANOVA, with Tukey's post hoc test (control, $n = 4$; ALS-*C9orf72*, $n = 5$; ISO-*C9orf72*, $n = 2$). $p < 0.05$ are annotated.

(legend continued on next page)

ALS.[35] Second, and perhaps most important, in contrast to chlordiazepoxide and alprazolam, acamprosate is not associated with sedation or respiratory depression, which are potentially hazardous side effects in the ALS population.[36,37]

Acamprosate is neuroprotective in a patient-derived *C9orf72* cell model

We assessed dose-response curves and potential levels of toxicity of acamprosate in motor neurons derived from induced pluripotent stem cells (iPSCs) obtained from two ALS patients carrying the *C9orf72* repeat expansion, two healthy individuals, and one isogenic control line (see Figure 6A for an outline of the workflow and Figure S6 for a representative immunohistochemistry staining of cell cultures). The drug did not exert a toxic effect on *C9orf72* or healthy motor neurons, even at the higher doses (30 μM for 72 h, viability measured using 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide assay; Figures S7 and S8). Using cleaved caspase-3, an established biochemical proxy for cell survival in *C9orf72*[38–40] and other genetic forms of ALS,[41] acamprosate strongly reduced cell death in *C9orf72*-derived motor neurons; the percentage of cleaved caspase-3^{+} cells was 1.8-fold lower at 10 μM, and 2.3-fold lower at 30 μM (half-maximal effective concentration = 0.271 μM; Figures 6B, 6C, and S9).

Based on this initial screening data, we selected the highest efficacious doses (10 and 30 μM) for additional testing. We confirmed the efficacy of acamprosate at protecting motor neurons in additional cell lines (five ALS patients carrying the *C9orf72* repeat expansion, four healthy individuals, and two isogenic control lines; Figure 6D). Again, the drug had minimal effects on healthy motor neurons and isogenic controls (Figures 6B and 6C).

Effect of acamprosate is comparable to riluzole, the current standard of care

Next, we compared the efficacy of acamprosate in preventing cell death to riluzole, the most widely prescribed drug for ALS, which extends life expectancy by an average of 3–6 months.[42] As expected, riluzole treatment reduced cell death in *C9orf72* motor neurons (1.4-fold decrease at 10 μM; Figure 6E). Notably, acamprosate exerted an average of 30% higher protection, either alone (Figure 6D) or combined with riluzole, than that observed with riluzole (Figure 6E).

DISCUSSION

We used our genetic and transcriptomic information to nominate medications that could be repurposed to delay symptom onset among *C9orf72* carriers, and our data support further investigation of acamprosate as a potential treatment for this common form of neurodegeneration. Acamprosate is a US Food and Drug Administration (FDA)-approved medication with a favorable safety profile that is prescribed to maintain alcohol abstinence. The drug modulates glutamate receptors,[43] a mechanism overlapping with riluzole, a commonly prescribed medication for ALS patients that slows clinical progression, at least in part, through its anti-glutamate properties.[42,44] Interestingly, acamprosate, alone or in combination with riluzole, exceeded the neuroprotective properties of riluzole in our cell-based model.

Our work supports previous research suggesting that a combination of baclofen and acamprosate may be a plausible therapeutic strategy for ALS; the combination prevented TDP-43 stress granule formation in a human osteosarcoma U2OS cell line overexpressing human TDP-43 and displayed neuroprotective effects in primary motor neurons derived from SOD1^{G93A} rat embryos.[35] However, our work nominating acamprosate was an independent effort based on large-scale human genomic data. It is also the first time that acamprosate has been nominated as a personalized treatment for *C9orf72*-related ALS-FTD. These studies offer evidence that acamprosate may benefit ALS patients, warranting further consideration for the development of acamprosate as a treatment within this population. Our cell-based assays also suggested that a combination of riluzole and acamprosate treatment is more effective than acamprosate alone. Such combination therapy has emerged as a promising approach in neurological diseases, perhaps reflecting the need to ameliorate multiple pathways to produce a clinical benefit.[45]

Drug repositioning has become an attractive option for drug discovery, as using de-risked compounds lowers development costs and shortens time lines.[46] Systematic approaches based on genetic and transcriptomic data integration are also promising.[47,48] An emerging theme in the pharmaceutical industry is that drugs targeting proteins and pathways with genetic evidence are more likely to succeed.[16,17] The effect of acamprosate on *C9orf72*-related disease may be due to its known glutaminergic and GABAergic properties, but this is currently unproven. Instead, our work reinforces the value of genomic data to rapidly nominate drugs for repurposing in situations where the underlying molecular mechanisms of disease are complex and where the mode of action of a drug is unclear. In support of this data-driven, mechanism-free approach, we highlight that acamprosate was first approved for use in alcohol-dependence syndrome 35 years ago, and yet precisely how the drug works in this condition remains speculative. This situation is hardly unique among CNS drugs. Despite decades of patient use, the exact molecular mechanism by which riluzole confers its survival advantage in patients with ALS is unknown. Delaying clinical trials until the mechanism of action of acamprosate is determined may be unnecessary, especially considering the current uncertainty about how the *C9orf72* repeat expansion causes disease and the lack of reliable animal models.

Our findings also partially explain the variable age at onset observed in *C9orf72* patients and confirm that genetic modifiers are crucial in determining the onset age in this patient population.[6,49] Individuals with *C9orf72*-related disease with a high genetic risk of sporadic ALS developed symptoms approximately

(E) The graph shows the percentage of cleaved caspase-3^{+} motor neurons derived from *C9orf72* patients treated with riluzole (10 μM) and riluzole plus acamprosate (10 and 30 μM). Dots represent the mean effect of each line. Data are mean ± SD. Comparisons were performed using a one-way ANOVA with Tukey's post hoc test ($n = 5$). $p < 0.05$ are annotated as follows: *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$; ****$p < 0.001$, *****$p < 0.0001$.
See also Figures S6–S9 and Table S10.

3 years earlier than patients with a low genetic risk. Interestingly, the genetic risk attributable to general ALS did not affect onset age among patients not carrying the *C9orf72* repeat expansion, as confirmed recently in a multicenter study.[50] Our findings support the hypothesis that ALS is a multistep process[51–54] and suggest that a personalized approach will be necessary to treat the various genetic forms of ALS and neurodegenerative disorders.

We identified ''neuronal cytoskeleton'' and ''axonal transport'' as the central pathways influencing the age at onset among *C9orf72* carriers. Interestingly, arginine-containing dipeptides generated from *C9orf72* repeat expansions are known to impede microtubule-based motility and axonal transport machinery in cells.[55] This pathogenic mechanism involves physical interactions of the dipeptides with the kinesin-containing motor complexes. Reinforcing this research, we found the rs113247976 variant within this gene to be among the most influential on onset age, with *C9orf72* patients carrying the *KIF5A* variant experiencing symptoms nearly 3.5 years earlier than patients who did not carry this variant (Table S7). These observations indicate that axonal transport is centrally involved in *C9orf72*-related neurodegeneration, and this cellular process may be a credible therapeutic target. Notwithstanding, our data also suggest that multiple cellular processes are disrupted by *C9orf72* mutation, underscoring the importance of our more comprehensive genomic approach to drug repurposing.

Previous research has investigated disease modifiers of *C9orf72*. However, these efforts were focused on individual loci and ignored the collective contribution of the ALS polygenic inheritance.[19,56–58] This knowledge gap was primarily due to the limited availability of large *C9orf72* datasets required for this type of analysis. In contrast, our approach was based on genome-wide research involving 1,516 ALS/FTD *C9orf72* carriers and 7,037 ALS non-carriers, making it one of the largest cohorts reported to date. We also focused on genomic variants outside the *C9orf72* locus, which are likely more amenable to therapeutic intervention with small molecules than the repeat expansion itself.

Our research underscores the critical contribution of common risk factors as modifiers of monogenic forms of ALS.[35] Unfortunately, data were unavailable to allow a similar analysis of other highly penetrant genetic mutations that cause ALS, such as *SOD1* and *TBK1*. However, the onset age among *SOD1* mutation carriers, such as the Ala5Val variant, is much more circumscribed than the variable age penetrance observed among *C9orf72* patients.[59] Nevertheless, our study provides a foundation for future research exploring how the common risk factors of ALS affect other monogenic forms of the disease and leveraging this information to repurpose medications.

In conclusion, we found that the risk of sporadic ALS is a critical determinant of the age at symptom onset among patients carrying the *C9orf72* repeat expansion. We identified various pathways and genes that play pivotal roles in the underlying cellular processes determining onset timing. We integrated our genomic data with transcriptomics to nominate drugs that could be repurposed to delay disease onset, an approach that has previously prioritized medications for chronic conditions such as obesity[60] and osteoporosis.[61] Acamprosate emerged as the primary contender from these analyses, and we subsequently confirmed that this oral, inexpensive medication reduces *C9orf72* motor neuron death in cell-based assays. Future research will explore how acamprosate may be exerting this effect. Nevertheless, our work pinpoints acamprosate for future exploration as a treatment to slow the manifestation of symptoms in this common genetic form of neurodegeneration. Our innovative and multidisciplinary approach could also fill the gap in therapeutic discovery in other complex neurodegenerative conditions, and we have made our computational pipeline publicly available to facilitate such work.

Limitations of the study

Our study has limitations. There was only a 3-year difference in onset age between the patients with the highest and the lowest genetic risk of sporadic ALS, representing a modest portion of the 40-year spread reported for *C9orf72* onset age. Despite this, the magnitude of this effect was comparable to the results reported for the *TMEM106B* locus in *C9orf72* patients.[19] Interestingly, the outcome we observed was independent of this locus. Our capacity to predict age at onset in this group of patients, as well as our ability to assess the contribution of ALS genetic risk on disease progression, will improve as larger datasets with comprehensive phenotypic data become available.

Another limitation of our study is that the underlying mechanism of the neuroprotective effect of acamprosate in *C9orf72*-derived motor neurons remains unknown. Although exploring such mechanisms was beyond the scope of this paper, future research should carefully address this gap, especially considering the implications it may have for the broader ALS patient population. Previously published data suggested that acamprosate may also have a neuroprotective effect on non-*C9orf72* ALS patients. A combination of acamprosate and baclofen (PXT864) has been proposed as a treatment for sporadic ALS. However, no clinical trial has been registered, and the company has discontinued its development, citing financial reasons.[62] We anticipate our work will lead to a renewed interest in the otherwise overlooked acamprosate as a treatment for this uniformly fatal neurodegenerative disease. Such repurposing of FDA-approved drugs for unmet medical needs has strategic advantages that shorten the development time line—preclinical safety testing can be shortened as the safety profile and pharmacokinetic profile of existing drugs are already established. The standard daily dose (1,998 mg) of acamprosate is known to cross the blood-brain barrier and alter brain physiology; patients with alcohol dependence who were treated with acamprosate showed highly significant suppression of glutamate levels in the anterior cingulate gyrus over 4 weeks.[63] These observations provide a quantitative biomarker of target engagement. It also supports the hypothesis that the protective effect observed in our cell-based assays at the relatively low dose of 1 μM will translate into human patients at the well-tolerated standard dose.

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Bryan Traynor (traynorb@mail.nih.gov).

Materials availability

This study did not generate new unique reagents.

Data and code availability

The summary statistics of the reference dataset genome-wide association study (GWAS) are publicly available at http://databrowser.projectmine.com/. The individual-level data for the training and test datasets are available on dbGaP (accession nos. phs000101.v5.p1 and phs001963.v1.p1). The data for the replication dataset were from a different study (principal investigator: Christopher Shaw, King's College London) and are available upon reasonable request. RNA-seq data from the motor cortex were obtained from the New York Genome Center, and RNA-seq data from the cerebellum were obtained from GEO (GEO: GSE67196). The figures were created using BioRender and Inkscape. The Sankey diagram was created using Visual Paradigm. The programming code used in this paper is available at https://github.com/sarasaezALS/C9orf72_AAO and https://zenodo.org/records/13259646 (https://doi.org/10.5281/zenodo.13259646) to facilitate the application of this methodology to other disorders. The individual-level polygenic risk scores generated for patients with the *C9orf72* mutation are available at https://zenodo.org/uploads/13769448. The connectivity scores for the drugs evaluated as treatments for *C9orf72*-related disease are available at https://zenodo.org/records/13769483.

CONSORTIA

The members of the International ALS Genomics Consortium are Robert H. Baloh, Robert Bowser, Christopher B. Brady, Alexis Brice, James Broach, William Camu, Ruth Chia, Adriano Chiò, John Cooper-Knock, Daniele Cusi, Jinhui Ding, Carsten Drepper, Vivian E. Drory, Travis L. Dunckley, Eva Feldman, Mary Kay Floeter, Pietro Fratta, Glenn Gerhard, J. Raphael Gibbs, Summer B. Gibson, Jonathan D. Glass, Stephen A. Goutman, John Hardy, Matthew B. Harms, Terry D. Heiman-Patterson, Lilja Jansson, Janine Kirby, Hannu Laaksovirta, John E. Landers, Francesco Landi, Isabelle Le Ber, Serge Lumbroso, Claire Guissart, Daniel JL. MacGowan, Nicholas J. Maragakis, Gabriele Mora, Kevin Mouzat, Liisa Myllykangas, Richard W. Orrell, Lyle W. Ostrow, Stuart Pickering-Brown, Erik P. Pioro, Stefan M. Pulst, John M. Ravits, Alan E. Renton, Wim Robberecht, Ekaterina Rogaeva, Jeffrey D. Rothstein, Erika Salvi, Sonja W. Scholz, Michael Sendtner, Pamela J. Shaw, Katie C. Sidle, Zachary Simmons, David J. Stone, Pentti J. Tienari, Bryan J. Traynor, John Q. Trojanowski, Juan C. Troncoso, Miko Valori, Philip Van Damme, Vivianna M. Van Deerlin, Ludo Van Den Bosch, and Lorne Zinman. The members of the ITALSGEN Consortium are Stefania M. Angelocola, Francesco P. Ausiello, Marco Barberis, Ilaria Bartolomei, Stefania Battistini, Enrica Bersano, Giulia Bisogni, Giuseppe Borghero, Maura Brunetti, Corrado Cabona, Andrea Calvo, Fabrizio Canale, Antonio Canosa, Teresa A. Cantisani, Margherita Capasso, Claudia Caponnetto, Patrizio Cardinali, Paola Carrera, Federico Casale, Adriano Chiò, Tiziana Colletti, Francesca L. Conforti, Amelia Conte, Elisa Conti, Massimo Corbo, Stefania Cuccu, Eleonora Dalla Bella, Eustachio D'Errico, Giovanni DeMarco, Raffaele Dubbioso, Carlo Ferrarese, Pilar M. Ferraro, Massimo Filippi, Nicola Fini, Gianluca Floris, Giuseppe Fuda, Salvatore Gallone, Giulia Gianferrari, Fabio Giannini, Maurizio Grassano, Lucia Greco, Barbara Iazzolino, Alessandro Introna, Vincenzo La Bella, Serena Lattante, Giuseppe Lauria, Rocco Liguori, Giancarlo Logroscino, Francesco O. Logullo, Christian Lunetta, Paola Mandich, Jessica Mandrioli, Umberto Manera, Fiore Manganelli, Giuseppe Marangi, Kalliopi Marinou, Maria Giovanna Marrosu, Ilaria Martinelli, Sonia Messina, Cristina Moglia, Maria Rosaria Monsurrò, Gabriele Mora, Lorena Mosca, Maria R. Murru, Paola Origone, Carla Passaniti, Cristina Petrelli, Antonio Petrucci, Angelo Pirisi, Susanna Pozzi, Maura Pugliatti, Angelo Quattrini, Claudia Ricci, Giulia Riolo, Nilo Riva, Massimo Russo, Mario Sabatelli, Paolina Salamone, Marco Salivetto, Fabrizio Salvi, Marialuisa Santarelli, Luca Sbaiz, Riccardo Sideri, Isabella Simone, Cecilia Simonini, Rossella Spataro, Raffaella Tanel, Gioacchino Tedeschi, Anna Ticca, Antonella Torriello, Stefania Tranquilli, Lucio Tremolizzo, Francesca Trojsi, Rosario Vasta, Veria Vacchiano, Giuseppe Vita, Paolo Volanti, Marcella Zollino, and Elisabetta Zucchi. The members of the SLAGEN Consortium are Vincenzo Silani, Isabella Fogh, Nicola Ticozzi, Antonia Ratti, Cinzia Tiloca, Silvia Peverelli, Cinzia Gellera, Giuseppe Lauria Pinter, Franco Taroni, Viviana Pensato, Barbara Castellotti, Giacomo P. Comi, Stefania Corti, Roberto Del Bo, Cristina Cereda, Mauro Ceroni, Stella Gagliardi, Lucia Corrado, Letizia Mazzini, Gianni Sorarù, Flavia Raggi, Gabriele Siciliano, Costanza Simoncini, Annalisa Lo Gerfo, Massimiliano Filosto, Maurizio Inghilleri, and Alessandra Ferlini. The members of Project MinE ALS Sequencing Consortium are Philip Van Damme, Philippe Corcia, Philippe Couratier, Patrick Vourc'h, Orla Hardiman, Russell McLaughlin, Marc Gotkine, Vivian Drory, Nicola Ticozzi, Vincenzo Silani, Jan H. van den Veldink, Leonard H. Berg, Mamede de Carvalho, Jesus S. Mora Pardina, Monica Povedano, Peter Andersen, Markus Weber, Ayşe Nazlı Başak, Ammar Al-Chalabi, Chris Shaw, Pamela J. Shaw, Karen E. Morrison, John E. Landers, and Jonathan D. Glass. The members of the American Genome Center are Adelani Adeleye, Camille Alba, Dagmar Bacikova, Clifton L. Dalgard, Daniel N. Hupalo, Elisa McGrath Martinez, Anthony R. Soltis, Gauthaman Sukumar, Coralie Viollet, and Matthew D. Wilkerson.

ACKNOWLEDGMENTS

This work was supported in part by the Intramural Research Program of the National Institutes of Health, the National Institute on Aging (ZIA-AG000933, ZIA-AG000934), the National Institute of Neurological Disorders and Stroke (NINDS), the National Center for Advancing Translational Sciences, and by Merck Sharp & Dohme Corp. (a subsidiary of Merck & Co., Kenilworth, NJ). B.J.T. received additional support from the Centers for Disease Control and Prevention, the Muscular Dystrophy Association, Microsoft Research, the Packard Center ALS Research at Johns Hopkins, and the ALS Association. L.F. is supported by the UK Medical Research Council (MR/V027735/1 and MR/V000470/1) and the Motor Neurone Disease Association (MND019801). A.N.B., C. Tunca, and E.B. gratefully acknowledge the use of the services and facilities of the Koç University Research Center for Translational Medicine (KUTTAM), funded by the Presidency of Turkey, Head of Strategy and Budget. R.S. is supported by the Barrow Neurological Foundation. This work was also generated within the European Reference Network for Neuromuscular Diseases (ERN EURO-NMD, https://ern-euro-nmd.eu). This study utilized the high-performance computational capabilities of the Biowulf Linux cluster at the National Institutes of Health. The authors would like to thank the Project MinE GWAS Consortium. This study also used genotype and clinical data from the Wellcome Trust Case Control Consortium and the HyperGenes Consortium. We thank the patients and research subjects who contributed samples for this study, Dr. Ruth Pfeiffer, Dr. Justin Kwan, the Laboratory of Neurogenetics staff, and the NINDS Neurodegeneration Clinic staff for their collegial support. The authors also acknowledge the Target ALS Human Post-mortem Tissue Core, the New York Genome Center

for Genomics of Neurodegenerative Disease, the ALS Association, and the TOW Foundation. Additional acknowledgments may be found in the supplemental information.

AUTHOR CONTRIBUTIONS

S.S.-A. and B.J.T. conceived and designed the study, acquired and analyzed the data, and wrote the paper. C.d.S.S., R. Chia, S.N.B., I.L., R.H., J.L., C.B., J.D., J.R.G., A.J., R.D., V.P., S.P., L.C., J.J.F.A.v.V., W.v.R., C. Tunca, E. Bayraktar, M.X., A.I., A.S., C. Tiloca, N.T., F.V., L.M., K.K., A.A.K., S.O.-M., F.R., M.F., S.C.P., A.P., S. Gagliardi, M.I., A.F., R.V., A. Calvo, C.M., A. Canosa, U.M., M.G., J.M., G.M., C.L., R.T., F.T., P.C., S. Gallone, M.B., D.G., M.S., C.F., E.S., G.P.C., S.C., R.D.B., M.C., G.L.P., F. Taroni, E.D.B., E.B., C.J.C., S.H.L., R. Chung, H.P., K.E.M., J.C.-K., P.J.S., G.B., R.J.B.D., C.L.D., S.W.S., A.A.-C., L.H.v.d.B., R.M., O.H., C.C., G. Sorarù, S.D., S. Chandran, S. Pal, A.R., C.G., K.J., T.D.-O., N.P., T.W., A.N., G. Siciliano, V.S., A.N.B., J.H.V., W.C., J.D.G., J.E.L., A. Chiò, R.S., C.E.S., L.F., and I.F. acquired and analyzed the data and edited the manuscript.

DECLARATION OF INTERESTS

B.J.T. holds patents on clinical testing and therapeutic intervention for the hexanucleotide repeat expansion of *C9orf72*.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
  - Human participants
  - Datasets
- METHOD DETAILS
  - SNP array-based genotyping data quality control procedures and imputation
  - Whole-genome sequencing of the test dataset
  - Genetic risk score generation and computation
  - Leave-one-out analysis and decile generation
  - Pathway analysis
  - Drug prioritization via gene-drug pattern matching
  - Drug prioritization via gene expression-pattern matching
  - Drug *in vitro* validation
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.xgen.2024.100679.

A video abstract is available at https://doi.org/10.1016/j.xgen.2024.100679#mmc2.

Received: April 1, 2024
Revised: July 2, 2024
Accepted: September 22, 2024
Published: October 21, 2024

REFERENCES

1. Larson, T.C., Kaye, W., Mehta, P., and Horton, D.K. (2018). Amyotrophic Lateral Sclerosis Mortality in the United States, 2011-2014. Neuroepidemiology *51*, 96–103. https://doi.org/10.1159/000488891.

2. Logroscino, G., and Piccininni, M. (2019). Amyotrophic Lateral Sclerosis Descriptive Epidemiology: The Origin of Geographic Difference. Neuroepidemiology *52*, 93–103. https://doi.org/10.1159/000493386.

3. Renton, A.E., Majounie, E., Waite, A., Simón-Sánchez, J., Rollinson, S., Gibbs, J.R., Schymick, J.C., Laaksovirta, H., van Swieten, J.C., Myllykangas, L., et al. (2011). A hexanucleotide repeat expansion in C9ORF72 is the cause of chromosome 9p21-linked ALS-FTD. Neuron *72*, 257–268. https://doi.org/10.1016/j.neuron.2011.09.010.

4. DeJesus-Hernandez, M., Mackenzie, I.R., Boeve, B.F., Boxer, A.L., Baker, M., Rutherford, N.J., Nicholson, A.M., Finch, N.A., Flynn, H., Adamson, J., et al. (2011). Expanded GGGGCC hexanucleotide repeat in noncoding region of C9ORF72 causes chromosome 9p-linked FTD and ALS. Neuron *72*, 245–256. https://doi.org/10.1016/j.neuron.2011.09.011.

5. Majounie, E., Renton, A.E., Mok, K., Dopper, E.G.P., Waite, A., Rollinson, S., Chiò, A., Restagno, G., Nicolaou, N., Simon-Sanchez, J., et al. (2012). Frequency of the C9orf72 hexanucleotide repeat expansion in patients with amyotrophic lateral sclerosis and frontotemporal dementia: a cross-sectional study. Lancet Neurol. *11*, 323–330. https://doi.org/10.1016/S1474-4422(12)70043-1.

6. Murphy, N.A., Arthur, K.C., Tienari, P.J., Houlden, H., Chiò, A., and Traynor, B.J. (2017). Age-related penetrance of the C9orf72 repeat expansion. Sci. Rep. *7*, 2116. https://doi.org/10.1038/s41598-017-02364-1.

7. Balendra, R., and Isaacs, A.M. (2018). C9orf72-mediated ALS and FTD: multiple pathways to disease. Nat. Rev. Neurol. *14*, 544–558. https://doi.org/10.1038/s41582-018-0047-2.

8. Sierksma, A., Escott-Price, V., and De Strooper, B. (2020). Translating genetic risk of Alzheimer's disease into mechanistic insight and drug targets. Science *370*, 61–66. https://doi.org/10.1126/science.abb8575.

9. Taymans, J.M., Fell, M., Greenamyre, T., Hirst, W.D., Mamais, A., Padmanabhan, S., Peter, I., Rideout, H., and Thaler, A. (2023). Perspective on the current state of the LRRK2 field. NPJ Parkinsons Dis. *9*, 104. https://doi.org/10.1038/s41531-023-00544-7.

10. Cacabelos, R. (2022). What have we learnt from past failures in Alzheimer's disease drug discovery? Expet Opin. Drug Discov. *17*, 309–323. https://doi.org/10.1080/17460441.2022.2033724.

11. Wilson, D.M., 3rd, Cookson, M.R., Van Den Bosch, L., Zetterberg, H., Holtzman, D.M., and Dewachter, I. (2023). Hallmarks of neurodegenerative diseases. Cell *186*, 693–714. https://doi.org/10.1016/j.cell.2022.12.032.

12. Mead, R.J., Shan, N., Reiser, H.J., Marshall, F., and Shaw, P.J. (2023). Amyotrophic lateral sclerosis: a neurodegenerative disorder poised for successful therapeutic translation. Nat. Rev. Drug Discov. *22*, 185–212. https://doi.org/10.1038/s41573-022-00612-2.

13. Gribkoff, V.K., and Kaczmarek, L.K. (2017). The need for new approaches in CNS drug discovery: Why drugs have failed, and what can be done to improve outcomes. Neuropharmacology *120*, 11–19. https://doi.org/10.1016/j.neuropharm.2016.03.021.

14. Lamb, J., Crawford, E.D., Peck, D., Modell, J.W., Blat, I.C., Wrobel, M.J., Lerner, J., Brunet, J.P., Subramanian, A., Ross, K.N., et al. (2006). The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease. Science *313*, 1929–1935. https://doi.org/10.1126/science.1132939.

15. Barretina, J., Caponigro, G., Stransky, N., Venkatesan, K., Margolin, A.A., Kim, S., Wilson, C.J., Lehár, J., Kryukov, G.V., Sonkin, D., et al. (2012). The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. Nature *483*, 603–607. https://doi.org/10.1038/nature11003.

16. Ochoa, D., Karim, M., Ghoussaini, M., Hulcoop, D.G., McDonagh, E.M., and Dunham, I. (2022). Human genetics evidence supports two-thirds of the 2021 FDA-approved drugs. Nat. Rev. Drug Discov. *21*, 551. https://doi.org/10.1038/d41573-022-00120-3.

17. Hingorani, A.D., Kuan, V., Finan, C., Kruger, F.A., Gaulton, A., Chopade, S., Sofat, R., MacAllister, R.J., Overington, J.P., Hemingway, H., et al. (2019). Improving the odds of drug development success through human genomics: modelling study. Sci. Rep. *9*, 18911. https://doi.org/10.1038/s41598-019-54849-w.

18. Chio, A., Borghero, G., Restagno, G., Mora, G., Drepper, C., Traynor, B.J., Sendtner, M., Brunetti, M., Ossola, I., Calvo, A., et al. (2012). Clinical characteristics of patients with familial amyotrophic lateral sclerosis carrying the pathogenic GGGGCC hexanucleotide repeat expansion of C9ORF72. Brain *135*, 784–793. https://doi.org/10.1093/brain/awr366.

19. Gallagher, M.D., Suh, E., Grossman, M., Elman, L., McCluskey, L., Van Swieten, J.C., Al-Sarraj, S., Neumann, M., Gelpi, E., Ghetti, B., et al. (2014). TMEM106B is a genetic modifier of frontotemporal lobar degeneration with C9orf72 hexanucleotide repeat expansions. Acta Neuropathol. *127*, 407–418. https://doi.org/10.1007/s00401-013-1239-x.

20. Goodrich, J.K., Singer-Berk, M., Son, R., Sveden, A., Wood, J., England, E., Cole, J.B., Weisburd, B., Watts, N., Caulkins, L., et al. (2021). Determinants of penetrance and variable expressivity in monogenic metabolic conditions across 77,184 exomes. Nat. Commun. *12*, 3505. https://doi.org/10.1038/s41467-021-23556-4.

21. Nalls, M.A., Escott-Price, V., Williams, N.M., Lubbe, S., Keller, M.F., Morris, H.R., and Singleton, A.B.; International Parkinson's Disease Genomics Consortium IPDGC (2015). Genetic risk and age in Parkinson's disease: Continuum not stratum. Mov. Disord. *30*, 850–854. https://doi.org/10.1002/mds.26192.

22. Saez-Atienzar, S., Bandres-Ciga, S., Langston, R.G., Kim, J.J., Choi, S.W., Reynolds, R.H., International ALS Genomics Consortium, ITALSGEN; Dewan, R., Abramzon, Y., Ahmed, S., et al. (2021). Genetic analysis of amyotrophic lateral sclerosis identifies contributing pathways and cell types. Sci. Adv. *7*, eabd9036. https://doi.org/10.1126/sciadv.abd9036.

23. Bandres-Ciga, S., Saez-Atienzar, S., Kim, J.J., Makarious, M.B., Faghri, F., Diez-Fairen, M., Iwaki, H., Leonard, H., Botia, J., Ryten, M., et al. (2020). Large-scale pathway specific polygenic risk and transcriptomic community network analysis identifies novel functional pathways in Parkinson disease. Acta Neuropathol. *140*, 341–358. https://doi.org/10.1007/s00401-020-02181-3.

24. Kim, N.Y., Lee, H., Kim, S., Kim, Y.J., Lee, H., Lee, J., Kwak, S.H., and Lee, S. (2024). The clinical relevance of a polygenic risk score for type 2 diabetes mellitus in the Korean population. Sci. Rep. *14*, 5749. https://doi.org/10.1038/s41598-024-55313-0.

25. Plym, A., Penney, K.L., Kalia, S., Kraft, P., Conti, D.V., Haiman, C., Mucci, L.A., and Kibel, A.S. (2022). Evaluation of a Multiethnic Polygenic Risk Score Model for Prostate Cancer. J. Natl. Cancer Inst. *114*, 771–774. https://doi.org/10.1093/jnci/djab058.

26. Sekimitsu, S., Xiang, D., Smith, S.L., Curran, K., Elze, T., Friedman, D.S., Foster, P.J., Luo, Y., Pasquale, L.R., Peto, T., et al. (2023). Deep Ocular Phenotyping Across Primary Open-Angle Glaucoma Genetic Burden. JAMA Ophthalmol. *141*, 891–899. https://doi.org/10.1001/jamaophthalmol.2023.3645.

27. Wang, F., Ghanzouri, I., Leeper, N.J., Tsao, P.S., and Ross, E.G. (2022). Development of a polygenic risk score to improve detection of peripheral artery disease. Vasc. Med. *27*, 219–227. https://doi.org/10.1177/1358863X211067564.

28. Van Mossevelde, S., van der Zee, J., Gijselinck, I., Sleegers, K., De Bleecker, J., Sieben, A., Vandenberghe, R., Van Langenhove, T., Baets, J., Deryck, O., et al. (2017). Clinical Evidence of Disease Anticipation in Families Segregating a C9orf72 Repeat Expansion. JAMA Neurol. *74*, 445–452. https://doi.org/10.1001/jamaneurol.2016.4847.

29. Kolberg, L., Raudvere, U., Kuzmin, I., Adler, P., Vilo, J., and Peterson, H. (2023). g:Profiler-interoperable web service for functional enrichment analysis and gene identifier mapping (2023 update). Nucleic Acids Res. *51*, W207–W212. https://doi.org/10.1093/nar/gkad347.

30. Lachmann, A., Schilder, B.M., Wojciechowicz, M.L., Torre, D., Kuleshov, M.V., Keenan, A.B., and Ma'ayan, A. (2019). Geneshot: search engine for ranking genes from arbitrary text queries. Nucleic Acids Res. *47*, W571–W577. https://doi.org/10.1093/nar/gkz393.

31. Sakaue, S., and Okada, Y. (2019). GREP: genome for REPositioning drugs. Bioinformatics *35*, 3821–3823. https://doi.org/10.1093/bioinformatics/btz166.

32. Lussier, Y.A., and Chen, J.L. (2011). The emergence of genome-based drug repositioning. Sci. Transl. Med. *3*, 96ps35. https://doi.org/10.1126/scitranslmed.3001512.

33. Subramanian, A., Narayan, R., Corsello, S.M., Peck, D.D., Natoli, T.E., Lu, X., Gould, J., Davis, J.F., Tubelli, A.A., Asiedu, J.K., et al. (2017). A Next Generation Connectivity Map: L1000 Platform and the First 1,000,000 Profiles. Cell *171*, 1437–1452.e17. https://doi.org/10.1016/j.cell.2017.10.049.

34. Kropiwnicki, E., Evangelista, J.E., Stein, D.J., Clarke, D.J.B., Lachmann, A., Kuleshov, M.V., Jeon, M., Jagodnik, K.M., and Ma'ayan, A. (2021). Drugmonizome and Drugmonizome-ML: integration and abstraction of small molecule attributes for drug enrichment analysis and machine learning. Database *2021*, baab017. https://doi.org/10.1093/database/baab017.

35. Boussicault, L., Laffaire, J., Schmitt, P., Rinaudo, P., Callizot, N., Nabirotchkin, S., Hajj, R., and Cohen, D. (2020). Combination of acamprosate and baclofen (PXT864) as a potential new therapy for amyotrophic lateral sclerosis. J. Neurosci. Res. *98*, 2435–2450. https://doi.org/10.1002/jnr.24714.

36. Plosker, G.L. (2015). Acamprosate: A Review of Its Use in Alcohol Dependence. Drugs *75*, 1255–1268. https://doi.org/10.1007/s40265-015-0423-9.

37. Tariq, A., Hill, N.S., Price, L.L., and Ismail, K. (2022). Incidence and Nature of Respiratory Events in Patients Undergoing Bronchoscopy Under Conscious Sedation. J. Bronchology Interv. Pulmonol. *29*, 283–289. https://doi.org/10.1097/LBR.0000000000000837.

38. Dafinca, R., Scaber, J., Ababneh, N., Lalic, T., Weir, G., Christian, H., Vowles, J., Douglas, A.G.L., Fletcher-Jones, A., Browne, C., et al. (2016). C9orf72 Hexanucleotide Expansions Are Associated with Altered Endoplasmic Reticulum Calcium Homeostasis and Stress Granule Formation in Induced Pluripotent Stem Cell-Derived Neurons from Patients with Amyotrophic Lateral Sclerosis and Frontotemporal Dementia. Stem Cell. *34*, 2063–2078. https://doi.org/10.1002/stem.2388.

39. Beckers, J., Tharkeshwar, A.K., Fumagalli, L., Contardo, M., Van Schoor, E., Fazal, R., Thal, D.R., Chandran, S., Mancuso, R., Van Den Bosch, L., and Van Damme, P. (2023). A toxic gain-of-function mechanism in C9orf72 ALS impairs the autophagy-lysosome pathway in neurons. Acta Neuropathol. Commun. *11*, 151. https://doi.org/10.1186/s40478-023-01648-0.

40. Zhang, Y.J., Jansen-West, K., Xu, Y.F., Gendron, T.F., Bieniek, K.F., Lin, W.L., Sasaguri, H., Caulfield, T., Hubbard, J., Daughrity, L., et al. (2014). Aggregation-prone c9FTD/ALS poly(GA) RAN-translated proteins cause neurotoxicity by inducing ER stress. Acta Neuropathol. *128*, 505–524. https://doi.org/10.1007/s00401-014-1336-5.

41. Hart, M.P., and Gitler, A.D. (2012). ALS-associated ataxin 2 polyQ expansions enhance stress-induced caspase 3 activation and increase TDP-43 pathological modifications. J. Neurosci. *32*, 9133–9142. https://doi.org/10.1523/JNEUROSCI.0996-12.2012.

42. Lacomblez, L., Bensimon, G., Leigh, P.N., Guillet, P., and Meininger, V. (1996). Dose-ranging study of riluzole in amyotrophic lateral sclerosis. Amyotrophic Lateral Sclerosis/Riluzole Study Group II. Lancet *347*, 1425–1431. https://doi.org/10.1016/s0140-6736(96)91680-3.

43. Kalk, N.J., and Lingford-Hughes, A.R. (2014). The clinical pharmacology of acamprosate. Br. J. Clin. Pharmacol. *77*, 315–323. https://doi.org/10.1111/bcp.12070.

44. Debono, M.W., Le Guern, J., Canton, T., Doble, A., and Pradier, L. (1993). Inhibition by riluzole of electrophysiological responses mediated by rat kainate and NMDA receptors expressed in Xenopus oocytes. Eur. J. Pharmacol. *235*, 283–289. https://doi.org/10.1016/0014-2999(93)90147-a.

45. Cummings, J.L., Tong, G., and Ballard, C. (2019). Treatment Combinations for Alzheimer's Disease: Current and Future Pharmacotherapy Options. J. Alzheimers Dis. *67*, 779–794. https://doi.org/10.3233/JAD-180766.

46. Pushpakom, S., Iorio, F., Eyers, P.A., Escott, K.J., Hopper, S., Wells, A., Doig, A., Guilliams, T., Latimer, J., McNamee, C., et al. (2019). Drug

repurposing: progress, challenges and recommendations. Nat. Rev. Drug Discov. *18*, 41–58. https://doi.org/10.1038/nrd.2018.168.

47. Jarada, T.N., Rokne, J.G., and Alhajj, R. (2020). A review of computational drug repositioning: strategies, approaches, opportunities, challenges, and directions. J. Cheminf. *12*, 46. https://doi.org/10.1186/s13321-020-00450-7.

48. Reay, W.R., and Cairns, M.J. (2021). Advancing the use of genome-wide association studies for drug repurposing. Nat. Rev. Genet. *22*, 658–671. https://doi.org/10.1038/s41576-021-00387-z.

49. Beck, J., Poulter, M., Hensman, D., Rohrer, J.D., Mahoney, C.J., Adamson, G., Campbell, T., Uphill, J., Borg, A., Fratta, P., et al. (2013). Large C9orf72 hexanucleotide repeat expansions are seen in multiple neurodegenerative syndromes and are more frequent than expected in the UK population. Am. J. Hum. Genet. *92*, 345–353. https://doi.org/10.1016/j.ajhg.2013.01.011.

50. van Rheenen, W., van der Spek, R.A.A., Bakker, M.K., van Vugt, J.J.F.A., Hop, P.J., Zwamborn, R.A.J., de Klein, N., Westra, H.J., Bakker, O.B., Deelen, P., et al. (2021). Common and rare variant association analyses in amyotrophic lateral sclerosis identify 15 risk loci with distinct genetic architectures and neuron-specific biology. Nat. Genet. *53*, 1636–1648. https://doi.org/10.1038/s41588-021-00973-1.

51. Al Khleifat, A., Iacoangeli, A., van Vugt, J.J.F.A., Bowles, H., Moisse, M., Zwamborn, R.A.J., van der Spek, R.A.A., Shatunov, A., Cooper-Knock, J., Topp, S., et al. (2022). Structural variation analysis of 6,500 whole genome sequences in amyotrophic lateral sclerosis. NPJ Genom. Med. *7*, 8. https://doi.org/10.1038/s41525-021-00267-9.

52. Al-Chalabi, A., Calvo, A., Chio, A., Colville, S., Ellis, C.M., Hardiman, O., Heverin, M., Howard, R.S., Huisman, M.H.B., Keren, N., et al. (2014). Analysis of amyotrophic lateral sclerosis as a multistep process: a population-based modelling study. Lancet Neurol. *13*, 1108–1113. https://doi.org/10.1016/S1474-4422(14)70219-4.

53. Chio, A., Mazzini, L., D'Alfonso, S., Corrado, L., Canosa, A., Moglia, C., Manera, U., Bersano, E., Brunetti, M., Barberis, M., et al. (2018). The multistep hypothesis of ALS revisited: The role of genetic mutations. Neurology *91*, e635–e642. https://doi.org/10.1212/WNL.0000000000005996.

54. Cady, J., Allred, P., Bali, T., Pestronk, A., Goate, A., Miller, T.M., Mitra, R.D., Ravits, J., Harms, M.B., and Baloh, R.H. (2015). Amyotrophic lateral sclerosis onset is influenced by the burden of rare variants in known amyotrophic lateral sclerosis genes. Ann. Neurol. *77*, 100–113. https://doi.org/10.1002/ana.24306.

55. Fumagalli, L., Young, F.L., Boeynaems, S., De Decker, M., Mehta, A.R., Swijsen, A., Fazal, R., Guo, W., Moisse, M., Beckers, J., et al. (2021). C9orf72-derived arginine-containing dipeptide repeats associate with axonal transport machinery and impede microtubule-based motility. Sci. Adv. *7*, eabg3013. https://doi.org/10.1126/sciadv.abg3013.

56. van Blitterswijk, M., Mullen, B., Wojtas, A., Heckman, M.G., Diehl, N.N., Baker, M.C., DeJesus-Hernandez, M., Brown, P.H., Murray, M.E., Hsiung, G.Y.R., et al. (2014). Genetic modifiers in carriers of repeat expansions in the C9ORF72 gene. Mol. Neurodegener. *9*, 38. https://doi.org/10.1186/1750-1326-9-38.

57. Zhang, M., Ferrari, R., Tartaglia, M.C., Keith, J., Surace, E.I., Wolf, U., Sato, C., Grinberg, M., Liang, Y., Xi, Z., et al. (2018). A C6orf10/LOC101929163 locus is associated with age of onset in C9orf72 carriers. Brain *141*, 2895–2907. https://doi.org/10.1093/brain/awy238.

58. Barbier, M., Camuzat, A., Hachimi, K.E., Guegan, J., Rinaldi, D., Lattante, S., Houot, M., Sánchez-Valle, R., Sabatelli, M., Antonell, A., et al. (2021). SLITRK2, an X-linked modifier of the age at onset in C9orf72 frontotemporal lobar degeneration. Brain *144*, 2798–2811. https://doi.org/10.1093/brain/awab171.

59. Bali, T., Self, W., Liu, J., Siddique, T., Wang, L.H., Bird, T.D., Ratti, E., Atassi, N., Boylan, K.B., Glass, J.D., et al. (2017). Defining SOD1 ALS natural history to guide therapeutic clinical trial design. J. Neurol. Neurosurg. Psychiatry *88*, 99–105. https://doi.org/10.1136/jnnp-2016-313521.

60. Liu, J., Lee, J., Salazar Hernandez, M.A., Mazitschek, R., and Ozcan, U. (2015). Treatment of obesity with celastrol. Cell *161*, 999–1011. https://doi.org/10.1016/j.cell.2015.05.011.

61. Brum, A.M., van de Peppel, J., van der Leije, C.S., Schreuders-Koedam, M., Eijken, M., van der Eerden, B.C.J., and van Leeuwen, J.P.T.M. (2015). Connectivity Map-based discovery of parbendazole reveals targetable human osteogenic pathway. Proc. Natl. Acad. Sci. USA *112*, 12711–12716. https://doi.org/10.1073/pnas.1501597112.

62. Pharnext (2023). In Pharnext refocuses its clinical trial programs on PXT3003, its most promising drug candidate, to optimize financial resources allocation, M.J. Elliott, ed.

63. Umhau, J.C., Momenan, R., Schwandt, M.L., Singley, E., Lifshitz, M., Doty, L., Adams, L.J., Vengeliene, V., Spanagel, R., Zhang, Y., et al. (2010). Effect of acamprosate on magnetic resonance spectroscopy measures of central glutamate in detoxified alcohol-dependent individuals: a randomized controlled experimental medicine study. Arch. Gen. Psychiatr. *67*, 1069–1077. https://doi.org/10.1001/archgenpsychiatry.2010.125.

64. Dewan, R., Chia, R., Ding, J., Hickman, R.A., Stein, T.D., Abramzon, Y., Ahmed, S., Sabir, M.S., Portley, M.K., Tucci, A., et al. (2021). Pathogenic Huntingtin Repeat Expansions in Patients with Frontotemporal Dementia and Amyotrophic Lateral Sclerosis. Neuron *109*, 448–460.e4. https://doi.org/10.1016/j.neuron.2020.11.005.

65. Nicolas, A., Kenna, K.P., Renton, A.E., Ticozzi, N., Faghri, F., Chia, R., Dominov, J.A., Kenna, B.J., Nalls, M.A., Keagle, P., et al. (2018). Genome-wide Analyses Identify KIF5A as a Novel ALS Gene. Neuron *97*, 1267–1288. https://doi.org/10.1016/j.neuron.2018.02.027.

66. Prudencio, M., Belzil, V.V., Batra, R., Ross, C.A., Gendron, T.F., Pregent, L.J., Murray, M.E., Overstreet, K.K., Piazza-Johnston, A.E., Desaro, P., et al. (2015). Distinct brain transcriptome profiles in C9orf72-associated and sporadic ALS. Nat. Neurosci. *18*, 1175–1182. https://doi.org/10.1038/nn.4065.

67. Choi, S.W., and O'Reilly, P.F. (2019). PRSice-2: Polygenic Risk Score software for biobank-scale data. GigaScience *8*, giz082. https://doi.org/10.1093/gigascience/giz082.

68. Duan, Y., Evans, D.S., Miller, R.A., Schork, N.J., Cummings, S.R., and Girke, T. (2020). signatureSearch: environment for gene expression signature searching and functional interpretation. Nucleic Acids Res. *48*, e124. https://doi.org/10.1093/nar/gkaa878.

69. Brooks, B.R. (1994). El Escorial World Federation of Neurology criteria for the diagnosis of amyotrophic lateral sclerosis. Subcommittee on Motor Neuron Diseases/Amyotrophic Lateral Sclerosis of the World Federation of Neurology Research Group on Neuromuscular Diseases and the El Escorial "Clinical limits of amyotrophic lateral sclerosis" workshop contributors. J. Neurol. Sci. *124*, 96–107. https://doi.org/10.1016/0022-510x(94)90191-0.

70. Neary, D., Snowden, J.S., Gustafson, L., Passant, U., Stuss, D., Black, S., Freedman, M., Kertesz, A., Robert, P.H., Albert, M., et al. (1998). Frontotemporal lobar degeneration: a consensus on clinical diagnostic criteria. Neurology *51*, 1546–1554. https://doi.org/10.1212/wnl.51.6.1546.

71. Salvi, E., Kutalik, Z., Glorioso, N., Benaglio, P., Frau, F., Kuznetsova, T., Arima, H., Hoggart, C., Tichet, J., Nikitin, Y.P., et al. (2012). Genomewide association study using a high-density single nucleotide polymorphism array and case-control design identifies a novel essential hypertension susceptibility locus in the promoter region of endothelial NO synthase. Hypertension *59*, 248–255. https://doi.org/10.1161/HYPERTENSIONAHA.111.181990.

72. van Rheenen, W., Shatunov, A., Dekker, A.M., McLaughlin, R.L., Diekstra, F.P., Pulit, S.L., van der Spek, R.A.A., Võsa, U., de Jong, S., Robinson, M.R., et al. (2016). Genome-wide association analyses identify new risk variants and the genetic architecture of amyotrophic lateral sclerosis. Nat. Genet. *48*, 1043–1048. https://doi.org/10.1038/ng.3622.

73. Das, S., Forer, L., Schönherr, S., Sidore, C., Locke, A.E., Kwong, A., Vrieze, S.I., Chew, E.Y., Levy, S., McGue, M., et al. (2016). Next-generation

genotype imputation service and methods. Nat. Genet. *48*, 1284–1287. https://doi.org/10.1038/ng.3656.

74. Chia, R., Sabir, M.S., Bandres-Ciga, S., Saez-Atienzar, S., Reynolds, R.H., Gustavsson, E., Walton, R.L., Ahmed, S., Viollet, C., Ding, J., et al. (2021). Genome sequencing analysis identifies new loci associated with Lewy body dementia and provides insights into its genetic architecture. Nat. Genet. *53*, 294–303. https://doi.org/10.1038/s41588-021-00785-3.
75. Chang, C.C., Chow, C.C., Tellier, L.C., Vattikuti, S., Purcell, S.M., and Lee, J.J. (2015). Second-generation PLINK: rising to the challenge of larger and richer datasets. GigaScience *4*, 7. https://doi.org/10.1186/s13742-015-0047-8.
76. Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A.R., Bender, D., Maller, J., Sklar, P., de Bakker, P.I.W., Daly, M.J., and Sham, P.C. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. Am. J. Hum. Genet. *81*, 559–575. https://doi.org/10.1086/519795.
77. Euesden, J., Lewis, C.M., and O'Reilly, P.F. (2015). PRSice: Polygenic Risk Score software. Bioinformatics *31*, 1466–1468. https://doi.org/10.1093/bioinformatics/btu848.
78. Mehta, P., Kaye, W., Raymond, J., Punjani, R., Larson, T., Cohen, J., Muravov, O., and Horton, K. (2018). Prevalence of Amyotrophic Lateral Sclerosis - United States, 2015. MMWR Morb. Mortal. Wkly. Rep. *67*, 1285–1289. https://doi.org/10.15585/mmwr.mm6746a1.
79. Leonenko, G., Baker, E., Stevenson-Hoare, J., Sierksma, A., Fiers, M., Williams, J., de Strooper, B., and Escott-Price, V. (2021). Identifying individuals with high risk of Alzheimer's disease using polygenic risk scores. Nat. Commun. *12*, 4506. https://doi.org/10.1038/s41467-021-24082-z.
80. Raudvere, U., Kolberg, L., Kuzmin, I., Arak, T., Adler, P., Peterson, H., and Vilo, J. (2019). g:Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). Nucleic Acids Res. *47*, W191–W198. https://doi.org/10.1093/nar/gkz369.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Anti-goat (donkey)	Thermo Fisher	Catalog # A21432, RRID: AB_2535853
Anti-guinea pig (goat)	Thermo Fisher	Catalog # A21450, RRID: AB_2535867
Anti-mouse (donkey)	Thermo Fisher	Catalog # A10037, RRID: AB_2534013
Anti-mouse (donkey)	Thermo Fisher	Catalog # A21202, RRID: AB_141607
Anti-rabbit (donkey)	Thermo Fisher	Catalog # A21206, RRID: AB_2535792
Anti-rabbit (donkey)	Thermo Fisher	Catalog # A10042, RRID: AB_2534017
Beta III tubulin (mouse)	Biolegend	Catalog # 801201, RRID: AB_2728521
Caspase-3 (rabbit)	Merck Millipore	Catalog # AB3623, RRID: AB_91556
ChAT (goat)	Merck Millipore	Catalog # AB144P, RRID: AB_2079751
Islet ½ (rabbit)	Abcam	Catalog # ab109517, RRID: AB_10866454
MAP-2 (guinea pig)	Synaptic systems	Catalog # 188004, RRID: AB_2138181
NeuN (mouse)	Merck Millipore	Catalog # MAB377, RRID: AB_2298772
Chemicals, peptides, and recombinant proteins		
Acamprosate calcium	Sigma-Aldrich	Catalog # A6981
Accutase	StemCell Technologies	Catalog # 07922
All-*trans* retinoic acid	Merck Millipore	Catalog # 554720
B27 supplement	Thermo Fisher	Catalog # 17504044
Brain Derived Neurotrophic Factor (BDNF)	Peprotech	Catalog # 450-02
Camptothecin (CPT)	Cell Signaling Technology	Catalog # 13637
CHIR 99021	Tocris Biosciences	Catalog # 4423
Ciliary Neurotrophic Factor (CNTF)	Peprotech	Catalog # 450-13
Compound E	Merck Millipore	Catalog # 530509
D(−)-2-Amino-5-phosphonopentanoic acid (D-AP5)	Sigma-Aldrich	Catalog # A8054
4,6-diamidino-2-phenylindole (DAPI)	Merck Millipore	Catalog # 508741
Dimethyl Sulfoxide (DMSO)	Merck Millipore	Catalog # 317275
Distilled H_2O	Merck Millipore	Catalog # EM3234
Dorsomorphin homolog 1 (DMH-1)	Tocris Biosciences	Catalog # 4126
GlutaMAX	Thermo Fisher	Catalog # 35050061
Hank's balanced salt solution (HBSS)	Thermo Fisher	Catalog # 14175
Insulin-like Growth Factor-I (IGF-1)	Peprotech	Catalog # 100-11
KnockOut DMEM/F-12	Thermo Fisher	Catalog # A1370801
mTeSR Plus Basal Medium	StemCell Technologies	Catalog # 100-0276
N2 supplement	Thermo Fisher	Catalog # 17502048

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Paraformaldehyde	Thermo Fisher	Catalog # 047392.9M
Penicillin/streptomycin	Merck Millipore	Catalog # 516106
Phosphate buffered saline (PBS)	Merck Millipore	Catalog # 6504
Purmorphamine (PMN)	Merck Millipore	Catalog # 540220
ReLeSR Passaging Reagent	StemCell Technologies	Catalog # 100-0484
Riluzole	Merck Millipore	Catalog # 557324
SB431542	Tocris Biosciences	Catalog # 1614
Triton X-100	Merck Millipore	Catalog # 648464
Vitronectin XF	StemCell Technologies	Catalog # 07180
Y27632 dihydrochloride	Tocris Biosciences	Catalog # 1254
Critical commercial assays		
MTT assay	Thermo Fisher	Catalog #V13154
SNP beadchip genotyping array	Illumina	Catalog #: InfiniumOmni2-5-8v1-4_A1
Deposited data		
ALS GWAS summary statistics	Van Rheenen et al.[50]	https://www.projectmine.com/research/download-data/
The individual-level data for the training and test datasets	Dewan et al.[64]	dbGaP (accession # phs001963.v1.p1), RRID: SCR_002709
The individual-level data for the training and test datasets	Nicolas et al.[65]	dbGaP (accession # phs000101.v5.p1), RRID: SCR_002709
The individual-level control data for the training and test datasets	https://dbgap.ncbi.nlm.nih.gov	dbGaP (accession # phs000001, phs000007, phs000187, phs000196, phs000292, phs000304, phs000315, phs000368, phs000372, phs000394, phs000397, phs000404, phs000421, phs000428, phs000615, phs000675, phs000801, and phs000869), RRID:SCR_002709
The individual-level data used as the replication dataset in this study.	King's College London, unpublished study	Available from study principal investigator (Christopher Shaw, King's College London) upon reasonable request.
The individual-level polygenic risk scores generated for the *C9orf72* patients.	This study	https://zenodo.org/uploads/13769448
LINCS dataset	Subramanian et al.[33]	https://lincsproject.org, RRID: SCR_006454
The LINCS connectivity scores for the drugs evaluated as treatments for *C9orf72*-related disease.	This study	https://zenodo.org/records/13769483
RNA-sequencing data from the motor cortex	New York Genome Center	https://www.nygenome.org
RNA-sequencing data from the cerebellum	Prudencio et al.[66]	GEO (accession # GSE67196), RRID: SCR_005012
Experimental models: Cell lines		
Human: iPS cell line	Cedars-Sinai	CS14iCTR-nxx
Human: iPS cell line	Coriell Biorepository	GM23338, RRID: CVCL_F182
Human: iPS cell line	University of Sheffield	MIFF1, RRID: CVCL_1E69
Human: iPS cell line	Cedars-Sinai	CS02iCTR-NTn1
Human: iPS cell line	University of Sheffield	ALS-183-C9
Human: iPS cell line	University of Sheffield	ALS-78
Human: iPS cell line	Cedars-Sinai	CS28iALS-C9nxx
Human: iPS cell line	Cedars-Sinai	CS29iALS-C9nxx
Human: iPS cell line	Cedars-Sinai	CS52iALS-C9nxx

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Human: iPS cell line	Cedars-Sinai	CS29iALS-C9n1.ISOxx
Human: iPS cell line	Cedars-Sinai	CS52iALS-C9n6.ISOxx
Software and algorithms		
Codes and scripts	This paper	https://github.com/sarasaezALS/C9orf72_AAO; https://zenodo.org/records/13259646
Drugmonizome	Kropiwnicki et al.[34]	https://maayanlab.cloud/drugmonizome/#/
ExperimentHub	https://www.bioconductor.org/packages/release/bioc/html/ExperimentHub.html	https://mrcieu.github.io/TwoSampleMR
Geneshot	Lachmann et al.[30]	https://maayanlab.cloud/genesgen, RRID: SCR_017582
g:Profiler2	Kolberg et al.[29]	https://biit.cs.ut.ee/gprofiler/ggos, RRID: SCR_018190
Harmony software	Perkin Elmer	Catalog # hh17000010, RRID: SCR_023543
PLINK (version 1.9)	PLINK Working Group	https://www.cog-genomics.org/plink/1.9/, RRID: SCR_001757
Prism (version 10)	GraphPad Software	https://www.graphpad.com, RRID: SCR_002798
PRSice2	Choi et al.[67]	https://choishingwan.github.ii/PRSice/, RRID: SCR_017057
Python	Python Team	https://www.python.org, RRID: SCR_008394
R (version 4.0.5)	R Core Team	https://www.r-project.org, RRID: SCR_001905
SignatureSearch (version 1.11.0)	Duan et al.[68]	https://bioconductor.org/packpack/release/bioc/vignettes/sigsignatureSea/inst/doc/signatursignatu.html, RRID: SCR_016177

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Human participants

Tables S8 and S9 list the source and clinical features of the cohorts used in this study. The ALS patients were diagnosed according to the El Escorial criteria[69] for the training and test datasets, and the FTD patients were analyzed according to the Neary criteria.[70] The *C9orf72* repeat expansions were detected using a repeat-primed polymerase chain reaction (PCR) assay according to an established protocol.[3] Written consent was obtained from all individuals enrolled in this study, and the institutional review board approved the study of the National Institute on Aging (protocol number 03-AG-N329). The *C9orf72* patients of the replication cohort were from a different study conducted at the King's College London and recruited by the SLAGEN Consortium, the University of Edinburgh, the Boğaziçi University, and Project MinE. Written consent was obtained from all individuals at their respective centers (see the *Replication cohort* section for details).

Datasets

Four independent datasets were used in the analysis, as is standard in genetic risk score analysis (see Figure 1 for the analysis workflow). The *reference dataset* consisted of summary statistics from a published GWAS based on 12,577 ALS cases and 23,475 control individuals (Table S9).[65] The allele weights obtained from the *reference dataset* were used to construct the ALS genetic risk score model in the training dataset.

The *training dataset* was composed of 7,037 ALS individuals known not to carry *C9orf72* repeat expansions and 34,235 controls (Table S9). The case samples were previously genotyped in the Laboratory of Neurogenetics, National Institutes of Health, using HumanOmniExpress SNP arrays (version 1.0, Illumina Inc., San Diego, CA).[22,65] The US control samples had been previously genotyped on Illumina SNP arrays as part of other GWAS efforts. These data were downloaded from the dbGaP repository (accession numbers phs000001, phs000007, phs000187, phs000196, phs000292, phs000304, phs000315, phs000368, phs000372, phs000394, phs000397, phs000404, phs000421, phs000428, phs000615, phs000675, phs000801, and phs000869). Additional

SNP array data from the *HYPERGENES* project and the Wellcome Trust Case Control Consortium were included as Italian and UK control subjects.[71] All study participants were of European ancestry, and familial cases were included in the analysis. The individual-level data for the *training dataset* was used as input for the *PRSice2* algorithm.

To assess the effect of the ALS genetic risk score on age at onset in *C9orf72* carriers, we calculate the genetic risk score in the *C9orf72* cohort. This *test dataset* consisted of 817 ALS/FTD cases that carry the *C9orf72* gene (see Table S9 for detailed information). Of these, 666 (81.5% of the cohort) were genotyped on SNP arrays in the Laboratory of Neurogenetics,[65] and 151 (18.4%) underwent whole-genome sequencing.[64] Genotype data for the 161 SNPs making up the model were extracted from the SNP array and whole-genome sequence data.

We replicated our findings in an independent *C9orf72* cohort obtained from a different study conducted at King's College London. This *replication dataset* consisted of 699 ALS/FTD cases known to carry the *C9orf72* repeat expansion and genotyped on SNP arrays using Illumina InfiniumOmni2-5-v1.[72]

METHOD DETAILS

SNP array-based genotyping data quality control procedures and imputation

Standard quality-control procedures were applied to the genotype data of the training dataset (n = 7,037 cases and 34,235 controls) and the test dataset (n = 666 cases) before input into the *PRSice2* algorithm. Briefly, individuals with low call rates (<95%), heterozygosity outliers (F-statistic > −0.15 or < 0.15), and ancestry outliers (+/− 6 standard deviations from means of principal components 1 and 2 of the 1000 Genomes phase 3 Caucasian with European ancestry from Utah (CEU) and Toscani in Siena, Italy (TSI) populations) were excluded. Variants with a missingness rate of >5%, exhibiting deviation from Hardy–Weinberg Equilibrium in controls (p-value < 10^{-6}), and palindromic SNPs were excluded. Cryptically related samples (defined as Pi_hat >0.125) were removed. The remaining sample genotypes were imputed using the Michigan Imputation Server pipeline using Minimac4[73] under default settings with Eagle (version 2.4) phasing based on *Haplotype Reference Consortium* (release 1.1 2016). Samples from the United States, Italy, the United Kingdom, Belgium, and France were imputed as a single group. Those variants were additionally filtered post-imputation to exclude variants with minor allele frequency <0.01, missing call rates >15%, and imputation quality R^2 < 0.3.

The SNP array-based genotyping in the *replication dataset* section describes the quality control and imputation methodology used in the replication dataset. Samples in common between the test and replication datasets were identified using the checksum program *id_geno_checksum.v2* and removed.

Whole-genome sequencing of the test dataset

In the test dataset, one hundred and fifty-one *C9orf72* carriers had previously undergone 150 base pair, paired-end whole-genome sequencing on a HiSeq X Ten sequencer.[64,74] Genotype data for the 161 SNPs making up the model were extracted from this whole-genome sequence data. They were merged with the SNP array genotype data for the remaining 666 *C9orf72* carriers using PLINK (version 1.9).[75,76]

Genetic risk score generation and computation

The genome-wide genetic risk score was calculated using the training dataset based on the weighted allele dosages obtained from the reference dataset as implemented in *PRSice2*.[67,77] This approach allows variants below the typical GWAS significance threshold of 5.0 x 10^{-8} to be included in the analysis, and the model selected a p-value threshold $\leq$ 0.0001 for SNP selection. For the training dataset, 1,000 permutations were used to generate empirical p-value estimates for each GWAS-derived p-value. Each permutation test in the training dataset provided a Nagelkerke's pseudo-R^2 value after adjusting for an estimated ALS prevalence of 5 per 100,000 of the population.[78] To avoid accounting for the *C9orf72* effect, 150 kb upstream and downstream of the *C9orf72* top GWAS variant, rs3849943 (chr9: 9:27,543,382; GRCh37), were removed from the analysis. Sex, age at onset, and principal components one to twenty were included as covariates in the model.

The *–score* command implemented in PLINK (version 1.9)[75] was used to test the general ALS genetic risk score's contribution to the age of symptom onset among the *C9orf72* expansion carriers. Risk allele dosages were counted, giving a dose of two if homozygous for the risk allele, one if heterozygous, and zero if homozygous for the reference allele. Linear regression was used to evaluate the association between the genetic risk scores and the age at onset, as implemented in *R* (version 4.0.3). Sex, disease diagnosis (ALS or FTD), and principal components one to twenty were included as covariates in the model. For replication, 147 out of 161 SNPs were used to build genetic risk scores. Sex and principal component 1 were used as covariates in the model. Genetic risk scores were transformed to Z scores based on cases (a Z score of one is equivalent to a single standard deviation of increase from the case population mean of the genetic risk score). There was a four-standard deviation difference from the genetic risk score mean between *C9orf72* carriers at the top and the bottom of the genetic risk score distribution. We calculated the maximum age at onset that the genetic risk score can explain by multiplying the slope of the regression (beta) by this standard deviation difference. This approach determined the extent to which the genetic risk score can account for variations in the age at onset. Individuals within the extremes were defined as those within the 3% tails of the genetic risk score distribution. On average, the score at the extremes exceeded ±2 standard deviations from the data mean (Table S2).[79]

Leave-one-out analysis and decile generation

Leave-one-out analyses were performed by iteratively excluding one variant from the ALS genetic risk score (based on 161 predictors) and re-estimating the causal effect on age at onset (Table S3). The 161 variants were then ordered based on the regression coefficient (beta) from the leave-one-out analysis and regrouped in ten deciles. Thus, we generated ten ranked deciles composed of 16 variants each. Decile ten contains the 16 variants with a more significant contribution to early age at onset. Regression analyses to evaluate the contribution of each decile (16 variants) to age at onset were performed using PLINK (version 1.9) as described previously.[75]

Pathway analysis

Functional enrichment analysis was performed using the g:GOSt function of g:Profiler2.[29] Briefly, SNPs from decile one and the *C9orf72* gene name were input for g:Profiler2.[29,80] Enrichment was performed against the biological process and molecular function from the Gene Ontology (GO) and the Kyoto Encyclopedia of Genes and Genomes (KEGG) databases. Only gene lists that contain between 5 and 500 genes were selected for the analysis. The significant threshold was an FDR-corrected *p*-value less than 0.05. Pathways containing a single gene were removed from the study.

Drug prioritization via gene-drug pattern matching

The SNPs that compose decile ten (Table S3) were mapped to gene names using the *g:SNPense* function of g:Profiler2.[29] These genes were used as search terms (defined as seed genes in Table S5) to identify functionally related genes based on previous knowledge and gene-gene co-expression data in the *Geneshot* webserver (accessed October 2021).[30] This application identifies genes associated with our search term based on their co-occurrence in publications and gene-gene similarity from human RNA-seq data (ARCHS4) to predict associations between genes and search terms.[30] The resulting list of genes plus the seed genes (Table S5) was then used as input for the GREP analysis (version 1.0.0), a pipeline identifying drugs that can be repurposed to target the gene set based on their enrichment in clinical indication categories.[31]

Finally, *drugmonizome* was used to identify significant biomedical terms within the specified drugs based on their indications (from the SIDER (Side Effect Resource) database, the mechanism of action (MOA, from the DrugRepurposingHub database), and the gene targets (from the Drugbank database) (data accessed 10/12/2022).[34]

Drug prioritization via gene expression-pattern matching

The nominated drugs were validated using the drug-repurposing algorithm called Connectivity map (*CMap*) through the *SignatureSearch* package using the LINCS search method (version 1.8.2).[68] This algorithm uses two inputs: (i) a disease gene expression signature based on a list of the up- and down-regulated genes, and (ii) a drug perturbation dataset composed of differential expression profiles of each gene after drug treatment. A bi-directional weighted Kolmogorov–Smirnov enrichment statistic test of gene expression ranks in the disease and drug signatures was used to assign a weighted normalized connectivity score (referred to as a *CMap* score) to each drug, reflecting the degree to which the drug 'flips' the signature of the disease. The *CMap* scores reflect the similarity between the query drug's signature and those in the database, suggesting that the drug has similar or dissimilar biological effects.

To build the *C9orf72* disease signature in the motor cortex, differentially expressed genes in *C9orf72* cases ($n = 36$) versus controls ($n = 58$) were selected at a Bonferroni-corrected *p*-value <0.05 (see supplemental information for details). This dataset was obtained from the New York Genome Center. To build the *C9orf72* disease signature in the cerebellum, differentially expressed genes in *C9orf72* cases ($n = 8$) versus controls ($n = 8$) were selected at FDR-corrected *p*-value <0.05. This dataset was downloaded as a raw count matrix from GEO (accession number GSE67196).[66] Drug signatures from the *LINCS* database were accessed through *ExperimentHub* (version 2.2.0) in the form of moderated z-scores from differential expression (DE) analysis of 12,328 genes from 8,140 compound treatments of 30 cell lines corresponding to a total of 45,956 signatures. Drug signatures corresponding to the nominated drugs were extracted from the analyzed results and were available for the following cell lines: A375 (LINCS ID = LCL-1235, a human cell line exhibiting epithelial morphology isolated from the skin of a patient with malignant melanoma, provided by the American Type Culture Collection (ATCC)), ASC (LCL-2104, human adipose stem cells, Sciencell Research Laboratories), FIBRNPC (LSC-1021, iPSC), HCC515 (LCL-2084, human cell line isolated from lung adenocarcinoma, Broad Institute), HT29 (LCL-1180, a human cell line with epithelial morphology isolated from a patient with colorectal adenocarcinoma, ATCC), NEU (LDC-1033, differentiated cell), A549 (LCL-1601, a human cell line isolated from the lung of a patient with non-small cell lung carcinoma, ATCC), NPC (LDC-1021, normal stem fibroblast-derived iPSCs), HA1E (LCL-2090, a human cell line isolated from kidney, Broad Institute), PC3 (LCL-1299, a human cell line isolated from a patient with grade IV prostate adenocarcinoma, ATCC), MCF7 (LCL-2138, a human cell line isolated from a patient with breast adenocarcinoma, ATCC), PHH (primary human hepatocyte), SKBR3 (LCL-1475, human cell line isolated from a patient with breast adenocarcinoma, ATCC), and VCAP (LCL-1147, a human cell line that was isolated from a patient with prostate carcinoma, ATCC). Additional information about the cell lines is available at lincsportal.ccs.miami.edu.

Drug *in vitro* validation

Human induced pluripotent stem cells (iPSCs) were maintained in 6-well plates coated with vitronectin XF (10 μg/mL) in the complete mTeSR-Plus medium. Media was replaced every 48 h, and cells were passaged as clumps every four to six days using ReLeSR,

according to the manufacturer's instructions. IPSCs were used between passages 20 and 35, and all iPSCs were cultured in 5% O_2 and 5% CO_2 at 37°C. In this study, we used iPS cells derived from four unaffected controls (CS14iCTR-21nxx, MIFF1, CS02iCTR-NTn1, GM23338), five iPS lines derived from ALS patients harboring *C9orf72* repeat expansions (ALS-183-C9, CS52iALS-C9nxx, ALS-78, CS28iALS-C9nxx, CS29iALS-C9nxx) and two isogenic control lines of CS52iALS-nxx and CS29iALS-C9nxx, respectively (CS52iALS-C9n6.ISOxx, CS29iALS-C9n1.ISOxx (*C9orf72* HRE Corrected)) (Table S10). The cells were fed on alternate days with the neuronal medium until day 40. Cells were previously characterized at day 40 of differentiation and were found to express classical mature motor neuron markers (ChAT, SMI32, Islet 1/2, MAP2, NeuN) (Figure S6 and Table S11).

Drug treatments were applied for 72 h at day 40 of differentiation, and cells were assayed for viability (MTT) or fixed for subsequent immunocytochemistry assays. All the imaging was performed using the Opera Phenix High Content Screening System (PerkinElmer) at ×40 magnification. The supplementary material provides a detailed description of the methodology.

QUANTIFICATION AND STATISTICAL ANALYSIS

The statistics and graphs for the cell line experiments were generated using GraphPad Prism 10 (GraphPad Software, La Jolla, California USA, www.graphpad.com). Comparisons were performed using one-way ANOVA or two-way ANOVA with Tukey's post hoc multiple comparisons test. Three technical replicates per treatment were averaged before plotting, and statistical analysis was performed using the percentage of cleaved caspase-3 mean across cell lines. Plots represent mean ± SD. *p*-values smaller than 0.05 are annotated unless otherwise specified in the figure legend.

Cell Genomics, Volume *4*

Supplemental information

Mechanism-free repurposing of drugs for *C9orf72*-related ALS/FTD using large-scale genomic data

Sara Saez-Atienzar, Cleide dos Santos Souza, Ruth Chia, Selina N. Beal, Ileana Lorenzini, Ruili Huang, Jennifer Levy, Camelia Burciu, Jinhui Ding, J. Raphael Gibbs, Ashley Jones, Ramita Dewan, Viviana Pensato, Silvia Peverelli, Lucia Corrado, Joke J.F.A. van Vugt, Wouter van Rheenen, Ceren Tunca, Elif Bayraktar, Menghang Xia, The International ALS Genomics Consortium, ITALSGEN Consortium, SLAGEN Consortium, Project MinE ALS Sequencing Consortium, Alfredo Iacoangeli, Aleksey Shatunov, Cinzia Tiloca, Nicola Ticozzi, Federico Verde, Letizia Mazzini, Kevin Kenna, Ahmad Al Khleifat, Sarah Opie-Martin, Flavia Raggi, Massimiliano Filosto, Stefano Cotti Piccinelli, Alessandro Padovani, Stella Gagliardi, Maurizio Inghilleri, Alessandra Ferlini, Rosario Vasta, Andrea Calvo, Cristina Moglia, Antonio Canosa, Umberto Manera, Maurizio Grassano, Jessica Mandrioli, Gabriele Mora, Christian Lunetta, Raffaella Tanel, Francesca Trojsi, Patrizio Cardinali, Salvatore Gallone, Maura Brunetti, Daniela Galimberti, Maria Serpente, Chiara Fenoglio, Elio Scarpini, Giacomo P. Comi, Stefania Corti, Roberto Del Bo, Mauro Ceroni, Giuseppe Lauria Pinter, Franco Taroni, Eleonora Dalla Bella, Enrica Bersano, Charles J. Curtis, Sang Hyuck Lee, Raymond Chung, Hamel Patel, Karen E. Morrison, Johnathan Cooper-Knock, Pamela J. Shaw, Gerome Breen, Richard J.B. Dobson, Clifton L. Dalgard, The American Genome Center, Sonja W. Scholz, Ammar Al-Chalabi, Leonard H. van den Berg, Russell McLaughlin, Orla Hardiman, Cristina Cereda, Gianni Sorarù, Sandra D'Alfonso, Siddharthan Chandran, Suvankar Pal, Antonia Ratti, Cinzia Gellera, Kory Johnson, Tara Doucet-O'Hare, Nicholas Pasternack, Tongguang Wang, Avindra Nath, Gabriele Siciliano, Vincenzo Silani, Ayşe Nazlı Başak, Jan H. Veldink, William Camu, Jonathan D. Glass, John E. Landers, Adriano Chiò, Rita Sattler, Christopher E. Shaw, Laura Ferraiuolo, Isabella Fogh, and Bryan J. Traynor

SUPPLEMENTARY METHODS

Replication cohort

The DNA samples for the replication cohort were obtained from a different study (principal investigator: Christopher Shaw, King's College London). These samples were collected at (A) King's College London and (B) Project MinE Sequencing Consortium (Utrecht University) as described below:

(A) The King's College London

The DNA samples of 464 novel *C9orf72* repeat expansion carriers were collected at King's College London. The participants were recruited through the SLAGEN Consortium, Boğaziçi University, and Scotland University. Details of the cohorts are as follows:

(1) **Italy (SLAGEN Consortium).** DNA samples of the novel Italian carriers were collected by the SLAGEN Consortium through the contribution of several Tertiary Centers and Clinical Laboratories across Italy. Patients were diagnosed with ALS according to the El Escorial revised criteria at the ALS tertiary referral center of Istituto Auxologico Italiano IRCCS [S1]. All patients had probable or definite familial ALS according to the Byrne criteria for FALS. Cognitive assessment in the care of ALS patients adopted standard neuropsychological assessment suitable for patients with verbal and motor impairment, such as the Edinburgh Cognitive and Behavioural ALS Screen (ECAS) [S2]. All individuals gave written informed consent, and the Ethics Committee of the Istituto Auxologico Italiano IRCCS, Milan, approved this protocol. Screening for the expanded repeats in the *C9orf72* gene was performed by a two-step protocol, including genotyping PCR followed by a repeat-primed PCR, as previously described [S3, S4].

(2) **Turkey (Boğaziçi University).** DNA samples of the Turkish carriers were collected at the Boğaziçi University and recruited across Turkey between 2002 and 2019. All individuals gave written informed consent, and the Ethics Committee on Research with Human Participants (INAREK) and Boğaziçi University, Istanbul, approved this protocol. Genomic DNA was isolated from whole blood using the MagNa Pure Compact System (Roche, Switzerland). The *C9orf72* GGGGCC repeat expansion was screened by Repeat-primed touchdown PCR using FastStart Universal Master Mix (Roche, Switzerland). FAM-labeled PCR products were subjected to fragment length analysis (Macrogen, Korea), and a saw-tooth pattern in expansion-positive cases was visualized in PeakScanner Software (ThermoFisher Scientific, USA) [S5, S6].

(3) **Scotland (Edinburgh University).** DNA samples were obtained from patients with ALS who donated blood for research to the Scottish Regenerative Neurology Tissue Bank as part of the Scottish Motor Neurone Disease (MND) Register. ALS patients were diagnosed following the 'El Escorial' criteria [S7]. Ethical approval for research analysis of the Scottish Regenerative Neurology Tissue Bank samples affiliated with the Scottish MND register was obtained from the East of Scotland Research Ethics Service.

(B) Project MinE Sequencing Consortium cohort (Utrecht University)

A total of 456 *C9orf72* carriers were obtained from the Project MinE data, which consisted of a collection of ALS patients recruited worldwide through the collaboration of tertiary referral clinics for motor neuron disease. Neurologists from the European participating Tertiary Centers who are members of the EU Joint Program – Neurodegenerative Diseases Research (JPND) project STRENGTH and the Project MinE Sequencing Consortium have agreed to follow shared standard parameters in the collection of clinical information of ALS patients. Neurologic examination and diagnostic tests were used to determine whether participants met the revised El Escorial criteria for possible, probable, laboratory-supported, or definite ALS as fully described elsewhere [S1, S8]. Details of the cohorts are as follows:

(1) **The Netherlands.** ALS patients were diagnosed with ALS at the tertiary referral clinic for motor neuron disease at the University Medical Center Utrecht (Dutch ALS Center) or were included in the Prospective ALS Study in The Netherlands. Patients were not pre-screened for any mutations related to ALS. All individuals gave written informed consent, and the University Medical Center Utrecht Medical Ethics Committee, Utrecht, approved this protocol.

(2) **UK MNDA Biobank.** Neurologists diagnosed cases with ALS in one of twenty UK hospitals specialized in motor neuron diseases, and patients had no family history of ALS. All participated in the UK National Biobank for Motor Neuron Disease Research. All individuals gave written informed consent, and the Trent University Medical Ethics Committee approved this protocol.

(3) **Turkey.** ALS patients were recruited from hospitals across Turkey between 2002 and 2019. DNA samples were collected at Boğaziçi University. A full description of individuals gave written informed consent, and the Ethics Committee on Research with Human Participants (INAREK) at Boğaziçi University, Istanbul, approved this protocol.

(4) **Belgium.** Patients were diagnosed with ALS at the tertiary referral clinic for motor neuron diseases at the University Hospitals in Leuven. All individuals gave written informed consent, and the Ethical Committee of the University Hospitals in Leuven approved this protocol.

(5) **Ireland.** Cases were diagnosed with probable or definite ALS according to the El Escorial Criteria by neurologists specialized in motor neuron diseases at Beaumont Hospital in Dublin [S7]. Patients were part of an ongoing population-based prospective ALS registry. Patients were selected for sequencing so that all areas of Ireland were adequately represented. All individuals reported Irish ancestry for at least three generations. All individuals gave written informed consent, and the Beaumont Hospital Research & Ethics Committee, Dublin, approved this protocol.

(6) **Spain.** According to the El Escorial criteria, ALS patients were diagnosed with definite or probable ALS [S7]. Neurologists and neurophysiologists saw patients at the tertiary referral centers: the Bellvitge Hospital and Carlos III Hospital for Catalonia and Madrid, respectively. All individuals gave written informed consent, and the Bellvitge University Hospital Ethics Committee, Barcelona, and "Comité de Ética de la Investigación del Hospital Carlos III," Madrid, approved this protocol.

(7) **The United States.** All samples were taken from patients seen at the Emory ALS Center in Atlanta, Georgia, USA. The Emory Center is a tertiary care ALS clinic that cares for many patients in Georgia and the surrounding states. Diagnoses were made by neurologists specializing in neuromuscular diseases and motor neuron diseases. After informed consent, complete demographic and clinical information was stored in the clinic database. DNA was collected and stored. All individuals gave written informed consent, and the Committee for the Protection of Human Subjects in Research of the University of Massachusetts Medical School, Worcester, approved this protocol.

(8) **France.** ALS patients were diagnosed with probable or definite ALS according to the El Escorial criteria by neurologists specialized in motor neuron diseases at the Reference centers for ALS of the University Hospitals of Limoges and Tours (LITORALS federation), members of the French FILSLAN networks [S7]. All individuals gave written informed consent, and the ethics committee of Tours Hospital and Limoges University Hospital approved this protocol.

(9) **Sweden.** Cases were diagnosed with probable or definite ALS according to the revised El Escorial Criteria by neurologists specialized in motor neuron diseases [S1]. All participants were of Swedish descent and had reported Northern Swedish citizenship for at least three generations. All individuals gave written informed consent, and the Regional Ethical Review Board in Umeå approved this protocol.

(10) **Israel.** ALS patients were diagnosed with probable or definite ALS according to the El Escorial criteria [S7] and in follow-up at the tertiary referral ALS clinic at the Hadassah University Hospital, Jerusalem, or Tel-Aviv Sourasky Medical Center in Tel-Aviv. Patients were not pre-screened for any

mutations related to ALS. Patients were referred from all regions in Israel and participated in a prospective ALS database and sample repository. All individuals gave written informed consent, and the Hadassah University Hospital Institutional Review Board, Hadassah, and The Institutional Review Board of Tel Aviv Sourasky Medical Center, Tel Aviv, approved this protocol.

(11) **Portugal.** According to the revised El-Escorial criteria [S1], neurologists specialized in motor neuron diseases diagnosed patients with possible, probable, or definite ALS. All individuals gave written informed consent, and the Local Research Ethics Committee at the Faculty of Medicine, University of Lisbon, approved this protocol.

(12) **Italy.** Patients were diagnosed with ALS according to the El Escorial revised criteria at the ALS tertiary referral center of Istituto Auxologico Italiano IRCCS [S1]. All patients had probable or definite familial ALS according to the Byrne criteria for FALS in the *SOD1*, *TARDBP*, *FUS*, and *C9orf72* genes. All individuals gave written informed consent, and the Ethics Committee of the IRCCS Istituto Auxologico Italiano, Milan, approved this protocol.

(13) **Switzerland.** ALS patients were diagnosed at the Muskelzentrum/ALS clinic at the Kantonsspital St. Gallen, a tertiary referral center in Northern Switzerland. Patients fulfilled the El-Escorial Criteria for probable lab supported, probable or definite, or ALS [S7]. All individuals gave written informed consent, and the Kantonale Ethikkomission des Kantons St. Gallen approved this protocol.

Overall, of the 836 ALS/FTD *C9orf72* repeat expansion carriers who passed the quality control (QC) thresholds, complete clinical information was available for 713 (n = 385 males and n = 328 females) individuals. Of those, only 699 did not overlap with the training dataset.

SNP array-based genotyping in the replication dataset

C9orf72 carriers (n=464) from the King's College London cohort were genotyped on the Illumina InfiniumOmni2-5-8v1-4_A1 platform in the Illumina certified laboratory of the Department of Social Genetic & Developmental Psychiatry, King's College London. Genotype raw data were first annotated to the dbSNP150 and merged after alignment to the same genomic coordinate (coordinates GRCh37). All multi-allelic and A/T or C/G SNPs were excluded. Pre-phasing quality control steps were performed according to PLINK's standard protocols (version 1.9) [S9]. SNPs were excluded by low call rate < 99% (--geno 0.01), minor allele frequency (MAF) > 0.01, and Hardy-Weinberg disequilibrium (HWE) < 1 x 10^{-6}. Individuals were removed by missingness genotype value of 3%, (--mind 0.03), by the +/-3SD to the mean of the inbreeding distribution F (+/- 0.25), if with mismatches between genetic and reported gender. Related and duplicated individuals were identified by calculating each pair of individuals' identity by state (IBS) status. Those who passed the threshold of PI_HAT > 0.175 were excluded from further analyses. Ancestry differences were estimated by principal components analysis (PCA) using EIGENSTRAT software, and the outliers identified by the first ten principal components (PCs) were removed.

After pre-imputation quality control, 836 individuals (457 males, 379 females) with coverage of 1,349,769 SNPs passed the filter thresholds. Filtered data were phased according to the Haplotype Reference Consortium Release 1.1 (HRC.r1-1) through the Wellcome Sanger Institute Imputation Service (https://www.sanger.ac.uk/tool/sanger-imputation-service/) adopting the Eagle method (version 2.4.1). Imputation analysis generated ~ 39,000,000 variants. Post-imputation QC was performed using QCTOOL (version 2.0.8) and SNPTEST (version 2.5.6) software. The quality of the inferred variants was estimated according to the following thresholds: INFO score > 0.6, average posterior probability (APP) > 0.9, MAF > 0.01, and Hardy-Weinberg disequilibrium > 1 x 10^{-6}. After post-imputation quality control, 7,635,605 SNPs remained for further analysis.

Whole-genome sequencing in the replication dataset

C9orf72 carriers (n=456) from Project MinE were whole-genome sequenced, and standard quality control criteria were applied. At the variant level, sites with a genotype quality (GQ) < 10 or missing and SNVs and indels with quality (QUAL) scores < 20 and < 30 were removed. Kinship coefficients (i.e., relatedness) were calculated using the KING method, as implemented in the *SNPRelate* package in R. All pairs of related individuals (kinship > 0.0625) were identified. The transition-transversion ratio in each sample was calculated using SnpSift27 (version 4.3p). The expected transition-transversion (Ti/Tv) ratio in whole-genome sequence data is ~2.0. Samples with a Ti/Tv ratio ± 6 SD from the entire distribution of samples were removed. The number of single nucleotide variants (SNV) and singletons was calculated per sample. Samples with total SNVs or singletons > 6 SD from the mean were removed.

The transition in sequencing platforms from HiSeq 2000 to HiSeq X caused an increase in observed indels per sample. Accordingly, samples were filtered by platform (HiSeq 2000 or HiSeq X) and were excluded if the number of indels was ± 6 SD from the mean of their respective group. After this step, the average sample depth was calculated again. It was higher for samples sequenced on the HiSeq 2000 (35X, on average) than for samples sequenced on the HiSeq X (25X, on average). However, no samples were removed at this step.

Samples with mismatched sex information or missing phenotypic information were excluded. The remaining sample quality control was performed on high-quality variants: multi-allelic SNVs, variants with missingness > 2%, variants with Hardy-Weinberg equilibrium p-value < 1×10^{-5}, variants with differential missingness between cases, and controls with p-value < 1×10 were removed. The final steps of sample quality control were performed on a set of variants with an MAF > 10%, SNP missingness < 0.1%, variants residing outside four complex regions (the major histocompatibility complex (MHC) on chromosome 6; the lactase locus (LCT) on chromosome 2; and inversions on chromosomes 8 and 17). A/T and C/G variants were also excluded. We used the SNVs to calculate observed and expected autosomal homozygous genotype counts for each sample, and they were removed if |F| > 0.1. Samples with a PI_HAT > 0.8 were excluded to avoid duplicated samples.

Principal component analysis, as implemented in EIGENSOFT, was used to visualize potential structure in the data induced by population stratification or other variables. Projections onto the HapMap3 and the 1000 Genomes (phase 3, version 5) populations indicated that the samples were primarily of European ancestry. However, some were of African or East Asian ancestries, while others appeared admixed. Outliers from the European population (HapMap3: > 10 SD on principal components (PC) 1-4, 1000 Genomes: > 4 SD on PCs 1-4) were excluded from further analyses.

All samples were sent in batches to Illumina for sequencing. Thus, all variants were regressed against batch using PLINK (version 1.9) [S9]. Finally, all variants with an association p-value < 1.0×10^{-10} in at least one batch were excluded.

New York Genome Center RNA sequencing

Patient and control samples were acquired from the New York Genome Center (NYGC) Consortium Database and can be accessed by contacting the NYGC at https://www.nygenome.org/contact/. *Trimmomatic* software was used to trim the original sequencing files obtained from NYGC to 80 base pairs to remove barcodes and improve sample quality [S10]. To account for sample sequencing depth differences, the reads from all samples were downsampled to 25 million reads. These reads were aligned to the hg38 reference genome (https://www.ncbi.nlm.nih.gov/assembly/GCF_000001405.26/) with the *Spliced Transcripts Alignment to a Reference* (STAR) software [S11].

These raw mRNA transcript counts were collated into a single count matrix file for differential expression analysis (DEA). First, samples in the count matrix file collected from the motor cortex were selected. Next, genes with a low number of transcript counts (5 or less) were removed, as well as those that were

significantly associated with biological sex (DESeq2 FDR adjusted p-value < 0.05). The NYGC provided information regarding which patients had a pathogenic C9orf72 expansion. The number of patients with a pathogenic *C9orf72* expansion only (i.e., no other known genetic predisposition to ALS) was 36. The DEA was performed on the resulting non-normalized count data matrix using the *DESeq2* package in R to compare the expression profiles of *C9orf72* ALS patients and controls [S12]. Differentially expressed genes (DEGs) were counted with a Bonferroni adjusted p-value < 0.05.

Individual-level variant analysis

We evaluated associations of the rs113247976 (*KIF5A*) variant with age at onset using linear regression models adjusted for sex and principal components one to ten. The variant was studied under an additive genotypic and a dominant genotypic model.

Induced pluripotent stem cells (iPSCs) maintenance and differentiation into motor neurons

Differentiation of iPSCs (**Table S10**) into motor neurons was performed as previously described with modifications [S13]. For the differentiation, induced pluripotent stem cells (iPSCs) were seeded in growth factor-reduced Matrigel-coated plates (0.1 µg/ml). On day zero, iPSCs at 100% confluence were washed once with PBS. Neuralization was initiated by switching to iPSC-NPC day 1-6 differentiation media (containing 50% KnockOut DMEM/F-12, 50% neurobasal medium, 0.5× N2 supplement, 0.5× B27 supplement, 1× GlutaMAX, 1% penicillin/streptomycin; this will be referred as a basal medium) supplemented with 2 µM dorsomorphin homolog 1 (DMH1), 2 µM SB431542, 3 µM CHIR (a GSK3 inhibitor), which was replaced every 24 hours. On day 7 of the differentiation, cells were switched to day 7-12 iPSC-NPC differentiation media (which contains basal medium, supplemented with 1 µM CHIR, 2 µM DMH1, 2 µM SB431542, 0.1 µM all-*trans* retinoic acid, and 0.5 µM purmorphamine (PMN)). For the passage, cells were washed with HBSS without calcium and magnesium and incubated for 7 minutes with Accutase at 37°C. Accutase was neutralized with double the medium quantity, and the cell suspension was centrifuged at 200 g for 4 minutes. The supernatant was discarded, and the cell pellet was resuspended in 7-12 iPSC-NPC differentiation media supplemented with 10µM Y27632 ROCK inhibitor. Cells were re-plated onto new matrigel-coated 6-well plates at a ratio of 1:1, and differentiation was continued. By day 12 of the differentiation, neural rosettes should have formed, and the cells should express classical neural progenitor cell (NPC) markers (Pax6, Nestin).

For the motor neuron differentiation, NPCs were plated in Matrigel-coated 6-well plates at a density of $7x10^5$ cells per well. After 24 hours of incubation, the medium was changed to basal medium supplemented with 0.5 µM all-*trans* retinoic acid and 0.1 µM PMN, and the medium was changed every day for six days. On day 19, the motor neuron progenitors were passaged with accutase onto matrigel-coated plates. The medium was replaced with basal medium supplemented with 0.5 µM all-*trans* retinoic acid, 0.1 µM PMN, 0.1 µM compound E (Cpd E), 10 ng/mL BDNF, 10 ng/mL CNTF, and 10 ng/mL IGF-1 (19-28 days medium) and seeded into an optic 96-well plate (Perkin Elmer) for staining, at a density of $2x10^4$ cells per well. On day 29 of the differentiation, the cells were switched to day 29-40 neuronal differentiation medium (which contains Neurobasal basal medium, supplemented with 1x of B27, 10 ng/mL BDNF, 10 ng/mL CNTF, and 10 ng/mL IGF-1). The cells were fed on alternate days with the neuronal medium until day 40. Cells were previously characterized at day 40 of differentiation and were found to express classical mature motor neuron markers (ChAT, SMI32, Islet 1/2, MAP2, NeuN) (**Figure S5**).

Drug Treatments

Acamprosate was obtained from Sigma-Aldrich (acamprosate calcium A6981), stored at 10 mM in dH_2O, and kept out of light at -20ºC until use. To evaluate the effect of acamprosate on neuronal survival, on day 40, motor neurons were treated with acamprosate (0.01-30 µM) diluted in 29-40 neuronal differentiation medium for 72 hours. As a positive control for cell death, motor neurons were treated with 2 µM

camptothecin (CPT) made up in day 29-40 neuronal differentiation medium for 1 hour at 37ºC. Control cultures were treated with dH_2O, the vehicle of dilution of acamprosate.

Apoptosis assessment

On day 40, motor neurons were treated with the selected drugs. After three days of treatment, motor neurons were fixed and stained for (i) active Caspase-3 to identify cells undergoing apoptosis and (ii) MAP2, a neuron-specific cytoskeletal protein, to define the cytoplasmic boundaries of cells. 4,6-diamidino-2-phenylindole (DAPI) was used for nuclear counterstain. Quantitative imaging analysis was conducted through the Opera Phenix high content Screening System at 40x magnification, using the Harmony software for analysis. The percentage of Caspase-3 positive cells and the number of fragmented nuclei were assessed per every condition. At least 25 fields were randomly selected and scanned per well of a 96-well plate in triplicate. To identify and remove any false readings generated by the system, three random treated and untreated wells were selected and counted manually (blind to the group).

Immunocytochemistry

Cells were washed with PBS and incubated with 4% paraformaldehyde warmed to approximately 37°C for 10 minutes at room temperature, then washed with PBS. Fixed motor neurons were permeabilized with 0.3% Triton X-100 in PBS for 5 minutes and incubated in 5% donkey serum blocking solution for 1 hour to block non-specific staining. Antibodies were diluted in 5% donkey serum. Cells were incubated with primary antibodies (**Table S11**) overnight at 4°C and washed three times with PBS (5 minutes per wash). Cells were incubated with AlexaFluor secondary antibodies (1:400 dilution) (**Table S11**) for 1 hour at room temperature in the dark. They were washed once with PBS before incubating in 1 µg/mL of DAPI for 5 minutes at room temperature in the dark. Cells were washed three times with PBS and stored in the dark at 4ºC until imaging.

Imaging

All the imaging was performed using the Opera Phenix™ High Content Screening System (Perkin Elmer) at × 40 magnification to allow high throughput analysis without experimental bias. Z-stacks of at least eight or more planes separated by 0.7 µm were obtained from a minimum of 25 fields per well from three technical replicate wells per experiment, thus assessing > 6000 cells per experiment. 405, 488, 594, and 647nm lasers, and the appropriate excitation and emission filters were used. Settings were kept consistent while taking images from all cultures. For active caspase-3 analysis, the total number of caspase-3 positive cells was counted using the automated image analysis software Harmony (Perkin Elmer) and divided by the total number of cells.

MTT assay

The effect of acamprosate on cell viability was assessed via 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyl-2H-tetrazolium bromide (MTT) assay. This colorimetric assay is based on the reduction of a yellow tetrazolium salt (3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide to purple formazan crystals by metabolically active cells. For this assay, motor neurons were seeded in clear 96-well plates, treated with different acamprosate concentrations, and incubated for 72 hours at 37°C. Motor neurons were treated with 2 µM camptothecin (CPT) as a positive control for cell death. After incubation, the media was removed, and the cells were washed with PBS (100µl per well). MTT solution was then added to a final concentration of 0.5 mg/mL and incubated at 37°C for 2 hours to allow the formation of formazan crystals. The cells were washed with PBS, and DMSO was added (100µl per well) to dissolve the formazan crystals. The plates were shaken for 10 minutes to lyse the cells. Plate absorbance was read

at 570nm in a PherAstar plate reader. The percentage of cell viability was normalized to the vehicle group.

SUPPLEMENTARY FIGURES

Figure S1. The sporadic ALS genetic risk influenced the age at onset among ALS *C9orf72* and FTD *C9orf72* carriers, related to Figure 3.

(A) The regression line shows the association between ALS genetic risk score and age at onset among 817 ALS/FTD patients carrying *C9orf72* repeat expansions (p-value = 0.024, beta = -0.765, 95% CI = -1.429 – -0.101) (ALS & FTD group), 666 ALS patients carrying *C9orf72* repeat expansions (p-value = 0.028, beta = -0.815, 95% CI = -1.54 – -0.09) (ALS group), and 151 FTD patients carrying *C9orf72* repeat expansions (p-value = 0.705, beta = -0.333, 95% CI = -2.06 – 1.39) (FTD group). The shadow areas represent the 90% confidence interval of the regression model. (B) The forest plot shows the meta-analysis results of the ALS and FTD groups (p-value 0.029, beta = -0.742, 95% CI = -1.412– -0.072).

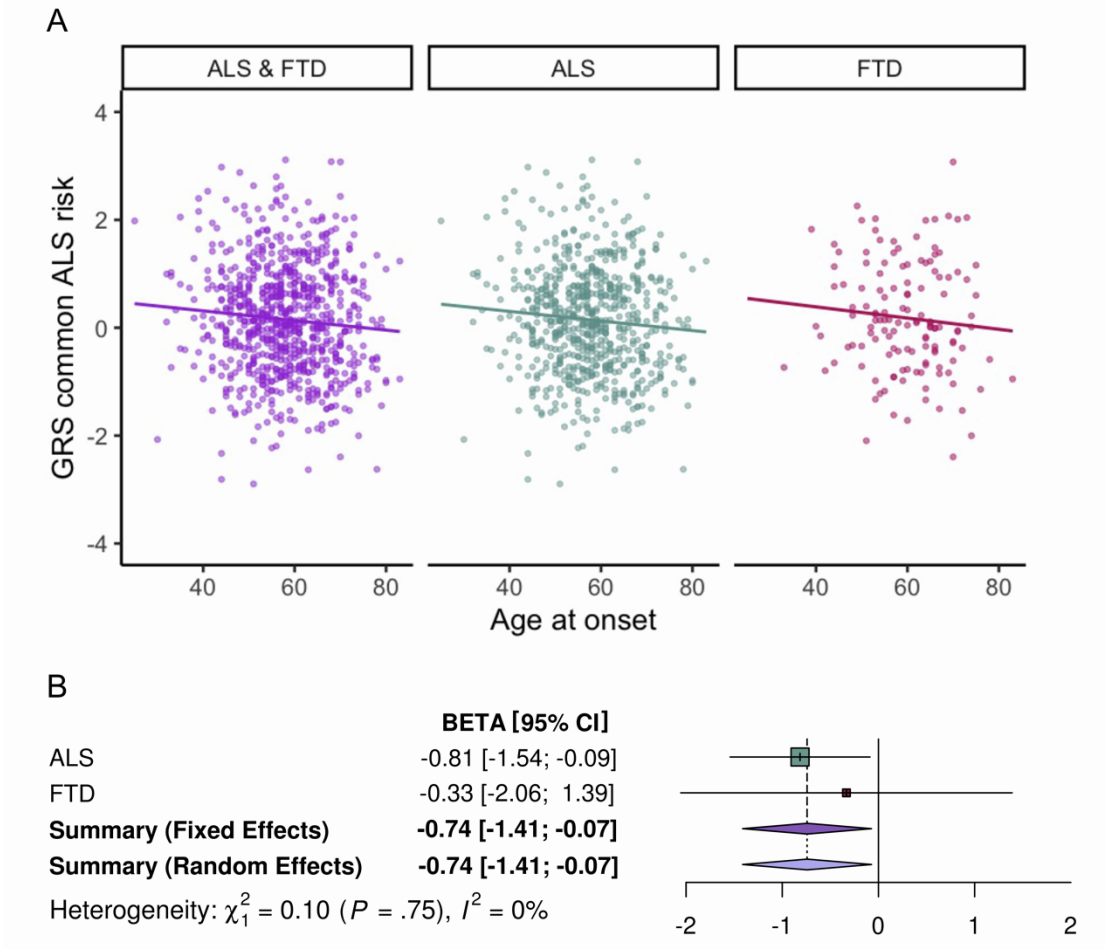


Figure S2. The sporadic ALS genetic risk score did not influence age at onset among patients without the *C9orf72* repeat expansion, related to Figure 3.

The regression line shows the lack of association between ALS genetic risk score and age at onset in 7,037 ALS patients without *C9orf72* repeat expansions (p-value = 0.437, beta = 0.115, 95% CI = -0.175-0.404). The shadow areas represent the 99% confidence interval of the regression model. The forest plot shows the regression result and the 95% confidence interval.

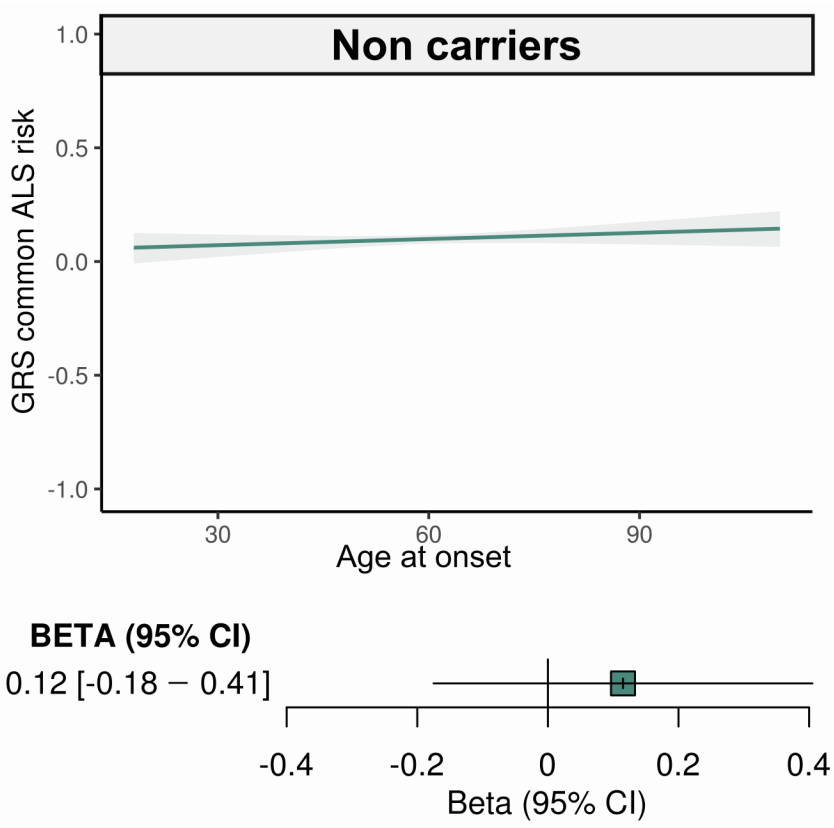


Figure S3. A schematic illustration of the leave-one-out analysis and decile calculation, related to Figure 4.

Leave-one-out analyses were performed by excluding one variant from the ALS genetic risk score analysis (based on 161 predictors) and re-estimating the causal effect on age at the onset of the remaining 160 variants. The variants were ordered based on their impact on age at onset, and ten ranked deciles, each containing 16 SNPs, were generated. Scores were recalculated using the deciles, and regression analysis evaluated the contribution of the 16 variants within each decile to age at onset.

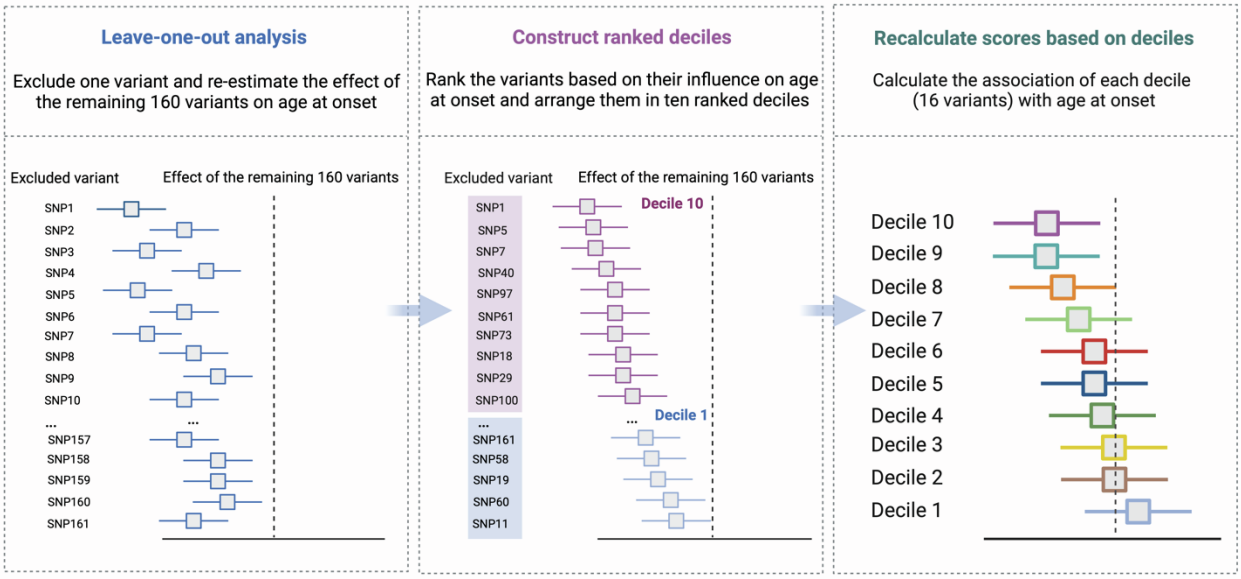


Figure S4. The genetic risk score of each decile showed no relationship to age of symptom onset among the ALS cases who did not carry the *C9orf72* expansion, related to Figure 4.

The Forest plot shows the effect estimates of the genetic risk score on age at onset in ALS non-*C9orf72* cases based on deciles obtained from the leave-one-out analysis. It also shows the regression results and the 95% confidence interval per decile. The analysis was performed in 7,037 ALS patients without *C9orf72* repeat expansions.

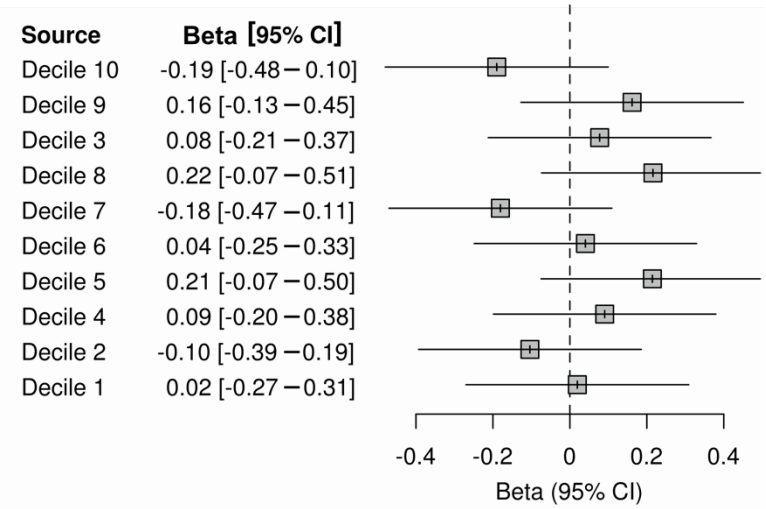


Figure S5. Enrichment analysis of the decile ten genes identified pathways influencing the age of onset among *C9orf72* carriers, related to Figure 4.

Twelve out of sixteen variants within decile ten are mapped to genes based on genomic coordinates. These genes were used for enrichment analysis based on Gene Ontology (GO) terms, including biological processes, molecular functions, and KEGG pathways.

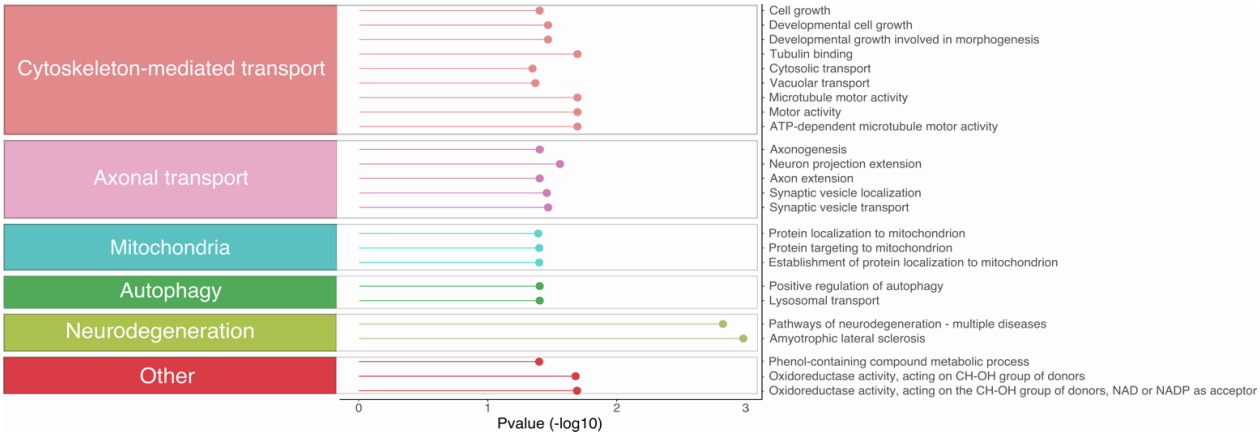


Figure S6. The cell lines displayed characteristics of motor neurons, related to Figure 6.

The picture shows motor neurons derived from unaffected controls (control), *C9orf72* ALS patients (ALS-C9orf72), and an isogenic control line (ISO-C9orf72). Chat, Choline acetyltransferase; DAPI, 4',6-diamidino-2-phenylindole; Islet 1/2, ISL LIM Homeobox 1/2; MAP2, Microtubule-associated protein 2; NeuN, neuronal nuclei antigen; SMI, neurofilament H; and TUJ, beta-tubulin III. The nuclei were counterstained with DAPI (Blue). Scale bar, 50 µm.

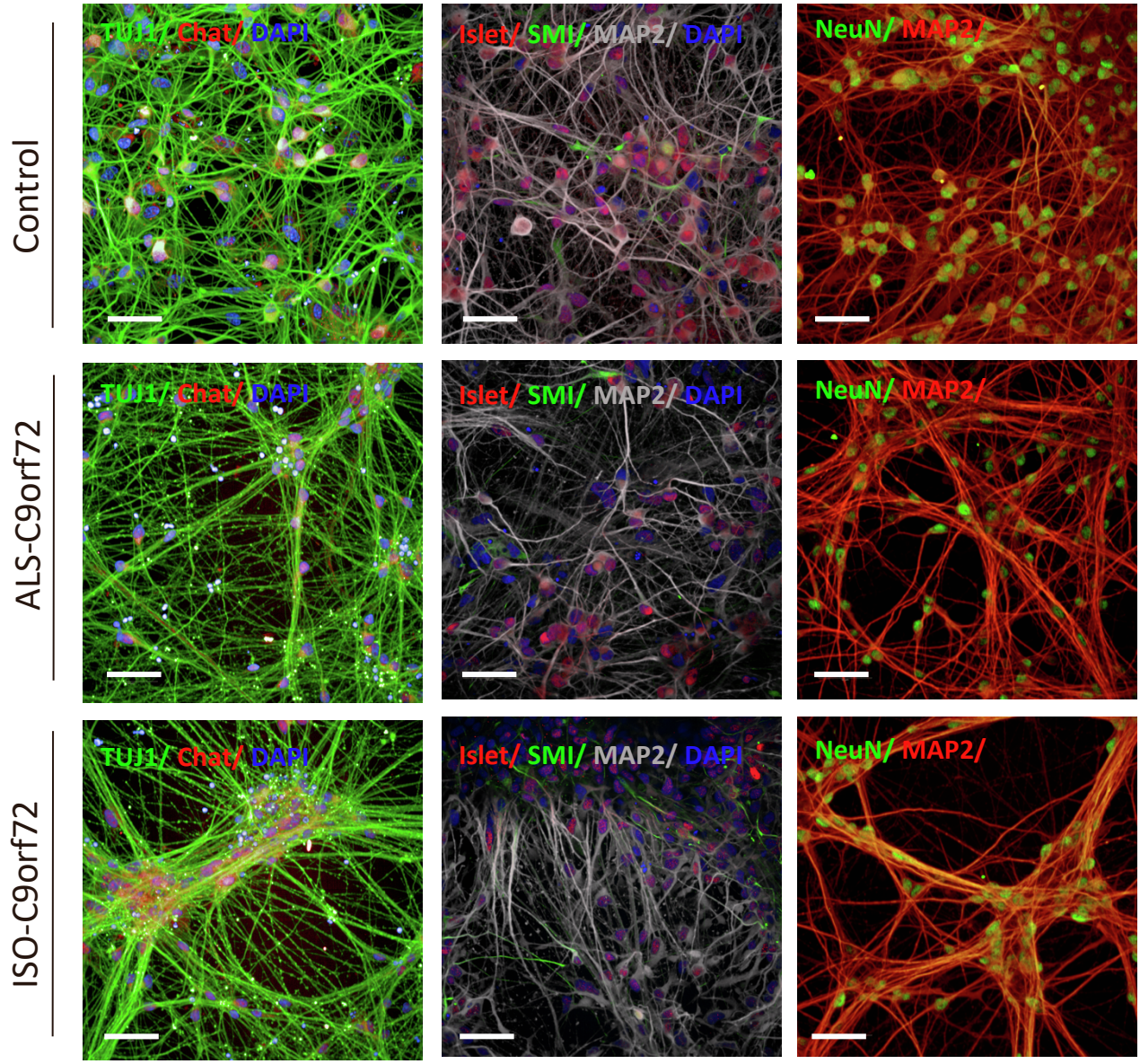


Figure S7. Acamprosate showed no evidence of toxicity in healthy-derived motor neurons, related to Figure 6.

Data are shown as the percentage of viable cells normalized to the vehicle (water). After 72 hours of treatment, acamprosate was not toxic for iPSC-derived motor neurons from two healthy donors. Camptothecin (CPT), an apoptosis inductor, was used as a positive control of cell death. The comparison was performed using one-way ANOVA. Each dot represents a technical replicate, and the data were pooled from two motor neuron lines to calculate the means and standard deviations.

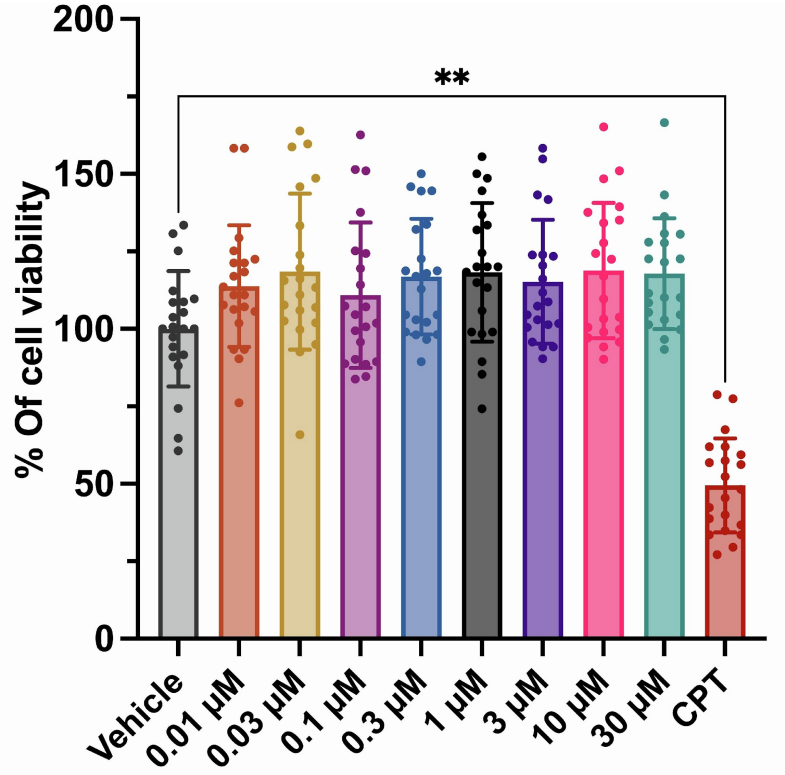


Figure S8. Acamprosate showed no evidence of toxicity in ALS-derived motor neurons, related to Figure 6.

Data are shown as the percentage of viable cells normalized to the vehicle (water). Acamprosate was not toxic in two iPSC-derived motor neuron cell lines derived from ALS patients carrying *C9orf72* expansion after 72 hours of treatment. The lines used were CS52iALS-C9nxx and CS28iALS-C9nxx; more information is listed in Table S10. Camptothecin (CPT), an apoptosis inductor, was used as a positive control of cell death. The comparison was performed using one-way ANOVA. Each dot represents a technical replicate, and the data were pooled to calculate the means and standard deviations.

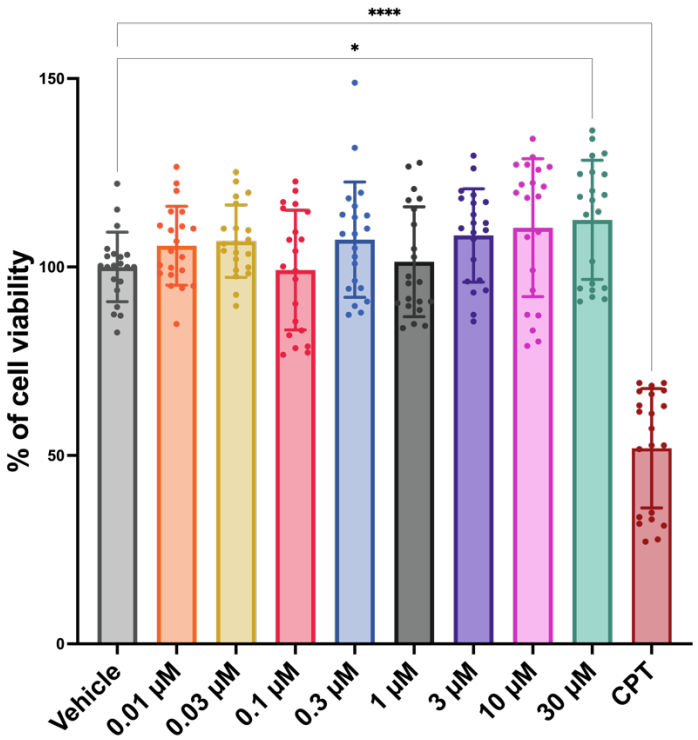
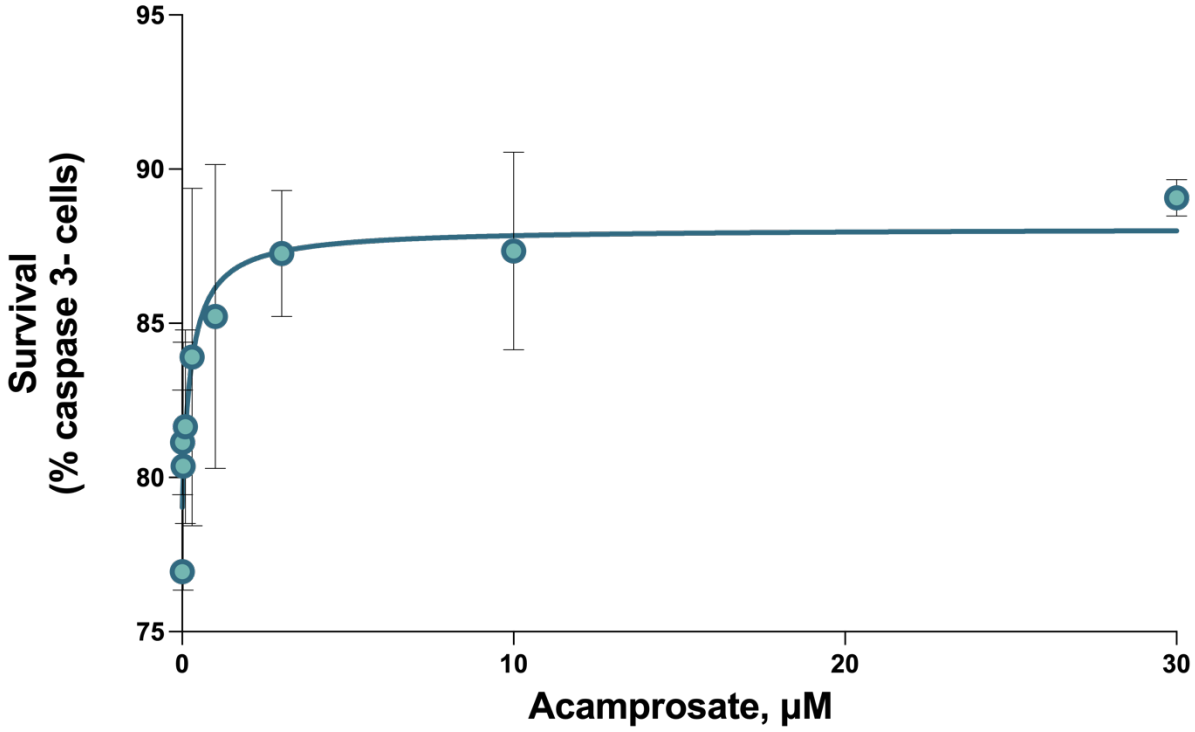


Figure S9. The dose-response curve of acamprosate in ALS-derived motor neurons showed a neuroprotective effect, related to Figure 6.

Dose-response curve and half maximal effective concentration (EC50 = 0.271 µM, 95% CI = 0.0218 – 2.95) of acamprosate in *C9orf72* ALS-derived motor neurons. The dose-response curve illustrates the effect of increasing concentrations of acamprosate (x-axis) on the survival of two different lines of *C9orf72*-carrying motor neurons (y-axis). For each line, data were averaged across three biological replicates, with three technical replicates for each one. Data are means ± standard deviations.



SUPPLEMENTARY TABLES

Table S1. The ALS genetic risk profile was based on 161 SNPs, related to Figure 2.

The polygenic risk score was generated using genetic data from Van Rheenen et al., 2016 [S14]. EA, effect allele; Chr, chromosome; Pos, position in build hg 38.

rsID	Chr	Pos (hg38)	EA	Gene names	Beta	P-value
rs10938692	4	8116834	T	*ABLIM2*	-0.079	$2.74x10^{-5}$
rs12369156	12	120729872	A	*ACADS*	0.261	$7.17x10^{-6}$
rs116488199	10	1675116	A	*ADARB2*	0.132	$8.83x10^{-5}$
rs1159918	4	99321852	A	*ADH1B*	-0.080	$2.03x10^{-5}$
rs320019	1	48610454	A	*AGBL4*	-0.078	$6.75x10^{-5}$
rs73103977	12	53513226	T	*ATF7*	-0.251	$1.78x10^{-5}$
rs11065961	12	111585263	A	*ATXN2*	-0.083	$5.81x10^{-5}$
rs6737916	2	32372917	A	*BIRC6*	0.121	$2.59x10^{-5}$
rs75087725	21	44333234	A	*CFAP410*	0.479	$8.65x10^{-11}$
rs10067826	5	10282407	A	*CMBL*	-0.143	$4.21x10^{-5}$
rs10443173	1	86068071	A	*COL24A1*	-0.106	$7.95x10^{-6}$
rs2271689	10	17046273	A	*CUBN*	-0.090	$9.43x10^{-5}$
rs6947666	7	137708989	A	*DGKI*	0.333	$4.92x10^{-5}$
rs10876069	12	50599395	T	*DIP2B*	0.072	$8.51x10^{-5}$
rs62073477	17	78448064	T	*DNAH17*	0.102	$4.65x10^{-5}$
rs77238283	17	11797907	T	*DNAH9*	0.167	$6.95x10^{-5}$
rs35059420	5	169995487	A	*DOCK2*	-0.159	$6.05x10^{-5}$
rs1442671	18	69528820	A	*DOK6*	0.078	$4.60x10^{-5}$
rs7764458	6	83116819	T	*DOP1A*	0.121	$5.74x10^{-5}$
rs11608027	11	34492865	T	*ELF5*	0.123	$3.05x10^{-5}$
rs17171046	7	37438260	T	*ELMO1*	0.116	$3.70x10^{-5}$
rs9901522	17	14770617	T	*ENSG00000205325*	0.146	$4.61x10^{-5}$
rs11185388	1	104198732	T	*ENSG00000215869*	-0.070	$9.25x10^{-5}$
rs2893656	7	106534655	A	*ENSG00000243797*	0.070	$8.87x10^{-5}$
rs3798105	5	133194937	T	*ENSG00000248245*	0.075	$9.76x10^{-5}$
rs4273590	5	159335610	A	*ENSG00000249738*	-0.191	$9.59x10^{-5}$
rs118072482	8	138017250	T	*ENSG00000253288*	0.297	$8.11x10^{-5}$
rs72973932	11	74400615	A	*ENSG00000254631*	0.182	$9.56x10^{-5}$
rs117219925	12	23142104	A	*ENSG00000256995*	-0.318	$7.15x10^{-5}$
rs111704832	15	93374070	T	*ENSG00000257060*	-0.155	$1.80x10^{-5}$
rs11171999	12	56846925	A	*ENSG00000258679*	0.077	$8.21x10^{-5}$
rs6603044	15	83015059	T	*ENSG00000259805*	-0.079	$1.08x10^{-5}$
rs56024498	16	76893238	A	*ENSG00000259995*	-0.070	$8.18x10^{-5}$
rs12991146	2	59884087	A	*ENSG00000271955*	-0.170	$6.05x10^{-5}$
rs144476584	9	23029711	T	*ENSG00000284418*	0.270	$7.46x10^{-5}$
rs116876275	13	65954214	T	*ENSG00000286395*	-0.287	$8.80x10^{-5}$

rs538622	5	172920676	A	*ERGIC1*	0.079	$1.33x10^{-5}$
rs2985994	13	45539849	T	*ERICH6B*	-0.081	$6.17x10^{-5}$
rs7930973	11	44159478	A	*EXT2*	0.076	$2.27x10^{-5}$
rs72792226	10	48204534	T	*FRMPD2*	0.285	$1.36x10^{-5}$
rs9903355	17	36580791	T	*GGNBP2*	-0.081	$5.29x10^{-6}$
rs7258235	19	2612120	A	*GNG7*	-0.156	$1.48x10^{-5}$
rs112820958	5	79465868	T	*HOMER1*	0.148	$4.75x10^{-5}$
rs11718653	3	122759011	T	*HSPBAP1*	0.221	$8.75x10^{-6}$
rs144049425	2	162505368	T	*KCNH7*	-0.251	$7.80x10^{-5}$
rs113247976	12	57581917	T	*KIF5A*	0.288	$1.13x10^{-5}$
rs61954176	13	40198035	T	*LINC00598*	-0.096	$5.32x10^{-6}$
rs11695294	2	176608890	A	*LINC01117*	0.108	$5.27x10^{-5}$
rs28407220	2	33816106	T	*LINC01320*	0.103	$7.11x10^{-6}$
rs150278778	1	209379025	A	*LINC01698*	-0.254	$4.82x10^{-5}$
rs72733862	5	8446181	A	*LINC02226*	0.148	$6.08x10^{-6}$
rs35318094	5	180245984	T	*MAPK9*	0.184	$1.15x10^{-5}$
rs17326496	5	113340882	T	*MCC*	-0.091	$5.54x10^{-5}$
rs79502718	18	50919413	A	*ME2*	0.169	$9.23x10^{-5}$
rs12972250	19	329746	A	*MIER2*	0.073	$6.10x10^{-5}$
rs9653747	21	18669100	A	*MIR548XHG*	-0.088	$4.64x10^{-6}$
rs12079484	1	181048307	A	*MR1*	0.075	$7.81x10^{-5}$
rs2240601	17	57673751	A	*MSI2*	-0.090	$5.07x10^{-5}$
rs4292737	8	10401605	A	*MSRA*	0.073	$5.11x10^{-5}$
rs4945276	11	78453938	T	*NARS2*	0.076	$5.90x10^{-5}$
rs150949995	5	150518937	T	*NDST1*	0.373	$4.28x10^{-6}$
rs642811	11	78053929	T	*NDUFC2-KCTD14*	0.085	$7.08x10^{-5}$
rs34432311	2	177346958	T	*NFE2L2*	-0.207	$3.20x10^{-5}$
rs68072647	17	9224590	A	*NTN1*	-0.077	$2.65x10^{-5}$
rs12886280	14	31829453	T	*NUBPL*	-0.083	$3.15x10^{-6}$
rs118036547	15	27863948	T	*OCA2*	0.251	$4.49x10^{-5}$
rs35346557	3	190120229	T	*P3H2*	-0.090	$4.92x10^{-5}$
rs3109207	4	168675360	A	*PALLD*	0.071	$9.49x10^{-5}$
rs36037136	1	164710739	A	*PBX1*	-0.291	$1.72x10^{-5}$
rs10492593	13	66919985	A	*PCDH9*	0.123	$2.89x10^{-5}$
rs2477866	1	233152025	A	*PCNX2*	-0.140	$3.02x10^{-5}$
rs16865645	2	177696567	T	*PDE11A*	0.096	$9.42x10^{-5}$
rs5766195	22	44921429	T	*PHF21B*	-0.072	$4.36x10^{-5}$
rs11652752	17	67379776	A	*PITPNC1*	-0.119	$6.49x10^{-5}$
rs8053191	16	81117558	T	*PKD1L2*	-0.152	$6.50x10^{-5}$
rs10430614	10	131935136	T	*PPP2R2D*	-0.079	$6.04x10^{-5}$
rs9355960	6	161901879	T	*PRKN*	-0.084	$1.37x10^{-5}$
rs2253050	16	74040378	T	*PSMD7-DT*	0.173	$7.43x10^{-5}$
rs28660489	12	64560784	A	*RASSF3*	-0.081	$6.13x10^{-5}$

rs12229321	12	64518074	T	*RASSF3*	0.071	$8.42x10^{-5}$
rs9813285	3	29329178	T	*RBMS3*	-0.077	$2.42x10^{-5}$
rs6683585	1	240967882	T	*RGS7*	-0.068	$9.16x10^{-5}$
rs143747467	16	11367549	T	*RMI2*	-0.293	$1.37x10^{-5}$
rs115348904	4	158390602	T	*RXFP1*	-0.208	$5.48x10^{-5}$
rs2294928	22	43986973	A	*SAMM50*	0.104	$4.27x10^{-5}$
rs35714695	17	28392769	A	*SARM1*	-0.134	$1.29x10^{-8}$
rs10139154	14	30678292	T	*SCFD1*	0.081	$1.92x10^{-5}$
rs118082508	12	56925035	T	*SDR9C7*	0.288	$3.76x10^{-5}$
rs111970477	3	47064091	A	*SETD2*	-0.262	$6.96x10^{-5}$
rs430979	4	2812971	T	*SH3BP2*	-0.085	$2.63x10^{-6}$
rs9995307	4	146488500	A	*SLC10A7*	0.080	$8.43x10^{-5}$
rs118038177	11	121459793	T	*SORL1*	-0.216	$4.79x10^{-5}$
rs60318796	17	32991387	T	*SPACA3*	0.109	$4.87x10^{-5}$
rs13387347	2	168898336	T	*SPC25*	0.075	$3.15x10^{-5}$
rs12967284	18	12532099	T	*SPIRE1*	0.079	$1.74x10^{-5}$
rs76805704	12	64138597	A	*SRGAP1*	0.188	$5.95x10^{-6}$
rs79612353	20	59884601	A	*SYCP2*	0.190	$3.84x10^{-5}$
rs112348322	4	118882809	A	*SYNPO2*	0.276	$2.60x10^{-5}$
rs74654358	12	64488187	A	*TBK1*	0.206	$7.72x10^{-7}$
rs11067262	12	114724621	T	*TBX3-AS1*	0.078	$8.56x10^{-5}$
rs79496463	8	132904843	T	*TG*	0.194	$7.17x10^{-5}$
rs13410191	2	137643025	A	*THSD7B*	-0.069	$9.98x10^{-5}$
rs651001	6	11569169	A	*TMEM170B*	-0.074	$3.81x10^{-5}$
rs115980385	7	141463227	T	*TMEM178B*	0.292	$9.81x10^{-5}$
rs10463311	5	151031274	T	*TNIP1*	-0.100	$8.51x10^{-7}$
rs4958888	5	151093281	A	*TNIP1*	0.089	$1.53x10^{-5}$
rs78549703	19	17638733	A	*UNC13A*	0.110	$1.31x10^{-8}$
rs8180839	7	5200339	A	*WIPI2*	0.151	$6.46x10^{-5}$
rs138116283	4	4318540	A	*ZBTB49*	0.337	$7.77x10^{-6}$
rs4974650	4	2309992	A	*ZFYVE28*	0.090	$1.63x10^{-6}$
rs8101883	19	56681170	A	*ZIM2-AS1*	0.084	$1.97x10^{-5}$
rs6997565	8	2560505	T	Intergenic	-0.146	$1.54x10^{-6}$
rs7118388	11	34432600	A	Intergenic	-0.084	$2.34x10^{-6}$
rs144387708	12	119264395	A	Intergenic	0.348	$4.95x10^{-6}$
rs116900480	12	58262322	T	Intergenic	0.294	$7.07x10^{-6}$
rs12900374	15	82741261	T	Intergenic	0.107	$7.27x10^{-6}$
rs10050775	5	38007940	A	Intergenic	-0.109	$1.08x10^{-5}$
rs117860708	11	1537217	A	Intergenic	0.237	$1.30x10^{-5}$
rs116946806	7	131997812	T	Intergenic	0.224	$1.43x10^{-5}$
rs4676496	3	39456514	A	Intergenic	0.077	$1.44x10^{-5}$
rs71472777	11	24121389	T	Intergenic	0.241	$1.45x10^{-5}$
rs112913348	5	108554187	T	Intergenic	-0.156	$1.51x10^{-5}$

rs34384833	5	91942338	A	Intergenic	0.265	$1.75x10^{-5}$
rs12472309	2	7266695	T	Intergenic	0.229	$2.45x10^{-5}$
rs10488631	7	128954129	T	Intergenic	-0.120	$2.52x10^{-5}$
rs7041171	9	111939350	T	Intergenic	-0.094	$2.71x10^{-5}$
rs12138742	1	119591406	T	Intergenic	-0.123	$2.74x10^{-5}$
rs62290425	4	4963737	A	Intergenic	0.126	$2.93x10^{-5}$
rs3098553	15	27631110	T	Intergenic	0.267	$3.04x10^{-5}$
rs79676202	12	49786775	T	Intergenic	0.204	$3.08x10^{-5}$
rs970258	2	5138399	T	Intergenic	-0.100	$3.11x10^{-5}$
rs72716562	5	7957484	A	Intergenic	-0.209	$3.12x10^{-5}$
rs141347161	7	42377714	T	Intergenic	0.149	$3.30x10^{-5}$
rs10008582	4	146026232	A	Intergenic	-0.106	$3.53x10^{-5}$
rs77058105	20	17070347	T	Intergenic	-0.198	$3.85x10^{-5}$
rs144129573	12	114485139	T	Intergenic	-0.209	$4.61x10^{-5}$
rs141730255	7	138253804	A	Intergenic	-0.180	$5.11x10^{-5}$
rs79446108	7	137863942	T	Intergenic	0.186	$5.33x10^{-5}$
rs6020200	20	50017227	A	Intergenic	-0.101	$5.36x10^{-5}$
rs6420358	13	84715333	A	Intergenic	-0.080	$6.14x10^{-5}$
rs12220832	10	80806689	T	Intergenic	0.123	$6.27x10^{-5}$
rs117452182	13	27293529	A	Intergenic	0.284	$6.31x10^{-5}$
rs7602576	2	112942040	T	Intergenic	-0.087	$6.39x10^{-5}$
rs11702120	21	23786163	A	Intergenic	0.202	$6.40x10^{-5}$
rs118071175	14	28515686	T	Intergenic	0.123	$6.78x10^{-5}$
rs7209200	17	5066645	T	Intergenic	-0.075	$6.82x10^{-5}$
rs9567838	13	47359837	T	Intergenic	0.095	$6.87x10^{-5}$
rs112288580	11	98729819	A	Intergenic	-0.290	$7.00x10^{-5}$
rs9819308	3	1691812	A	Intergenic	-0.086	$7.06x10^{-5}$
rs6037557	20	405220	T	Intergenic	-0.097	$7.56x10^{-5}$
rs9956309	18	38136697	A	Intergenic	-0.091	$8.42x10^{-5}$
rs76323495	16	16859460	A	Intergenic	-0.200	$8.54x10^{-5}$
rs16905848	11	20223160	T	Intergenic	-0.097	$8.72x10^{-5}$
rs35851984	17	28239061	A	Intergenic	-0.070	$8.98x10^{-5}$
rs76427181	6	86684739	A	Intergenic	-0.224	$9.38x10^{-5}$
rs2176039	22	45189151	A	Intergenic	0.069	$9.44x10^{-5}$
rs1570281	6	146517546	A	Intergenic	-0.077	$9.66x10^{-5}$
rs72838433	2	127913724	A	Intergenic	0.073	$9.87x10^{-5}$
rs193044924	17	15767328	A	Intergenic	-0.077	$9.89x10^{-5}$
rs73152707	3	86619077	T	Intergenic	0.170	$9.94x10^{-5}$
rs1146342	1	118440554	T	Intergenic	0.269	$9.98x10^{-5}$

Table S2. The genetic risk score influenced the age of symptom onset among *C9orf72* carriers, related to Figure 3.

The *bottom 3%* group is composed of individuals whose Z-score is in the bottom 3% of the genetic risk score distribution among *C9orf72* carriers and non-carriers, respectively. The *medium* group is comprised of individuals whose z-score is between 20-80% of the genetic risk score distribution. The *top 3%* group is comprised of individuals whose z-score is between 97-100% of the genetic risk score distribution.

	Age at onset mean	Age at onset standard deviation	Z-score mean	Z-score standard deviation	Count
Carriers					
Bottom 3%	58.11	10.23	-2.21	0.4	27
Medium	57.51	9.25	0.04	0.49	488
Top 3%	55.12	9.92	2.35	0.36	32
Non-Carriers					
Bottom 3%	60.25	12.18	-2.2	0.35	213
Medium	60.07	12.48	-0.03	0.47	4288
Top 3%	60.07	12.63	2.21	0.4	280

Table S3. The leave-one-out analysis stratified the 161 SNPs of the ALS genetic risk into ten deciles, related to Figure 4.

Decile ten contains the variants with a more significant contribution to age at onset. LOO, leave-one-out (indicating the removed SNP); Chr, chromosome; Pos (hg38), genomic position, GRCh38 assembly; SE, standard error.

LOO	Chr	Pos (hg38)	Gene name	Beta	SE	P-value	Decile	Rank
rs9901522	17	14770617	*ENSG00000205325*	-0.678	0.336	0.044	10	1st
rs118036547	15	27863948	*OCA2*	-0.705	0.337	0.037	10	2nd
rs2294928	22	43986973	*SAMM50*	-0.708	0.336	0.036	10	3rd
rs61954176	13	40198035	*LINC00598*	-0.711	0.337	0.035	10	4th
rs62073477	17	78448064	*DNAH17*	-0.715	0.337	0.034	10	6th
rs79502718	18	50919413	*ME2*	-0.715	0.336	0.034	10	5th
rs113247976	12	57581917	*KIF5A*	-0.716	0.336	0.034	10	7th
rs118082508	12	56925035	*SDR9C7*	-0.716	0.337	0.034	10	8th
rs1146342	1	118440554	Intergenic	-0.718	0.336	0.033	10	9th
rs117219925	12	23142104	*ENSG00000256995*	-0.72	0.337	0.033	10	10th
rs116900480	12	58262322	*Intergenic*	-0.721	0.336	0.032	10	11th
rs10876069	12	50599395	*DIP2B*	-0.725	0.336	0.031	10	12th
rs6420358	13	84715333	Intergenic	-0.727	0.336	0.031	10	13th
rs7118388	11	34432600	Intergenic	-0.728	0.336	0.031	10	14th
rs9355960	6	161901879	*PRKN*	-0.728	0.336	0.031	10	15th
rs10938692	4	8116834	*ABLIM2*	-0.73	0.336	0.03	10	16th
rs10488631	7	128954129	Intergenic	-0.731	0.337	0.03	9	17th
rs34384833	5	91942338	Intergenic	-0.732	0.336	0.03	9	18th
rs116876275	13	65954214	*ENSG00000286395*	-0.733	0.336	0.03	9	19th
rs12972250	19	329746	*MIER2*	-0.738	0.336	0.029	9	20th
rs11608027	11	34492865	*ELF5*	-0.739	0.336	0.028	9	22nd
rs17171046	7	37438260	*ELMO1*	-0.739	0.336	0.028	9	24th
rs62290425	4	4963737	Intergenic	-0.739	0.336	0.028	9	21st
rs6683585	1	240967882	*RGS7*	-0.739	0.337	0.028	9	23rd
rs6737916	2	32372917	*BIRC6*	-0.739	0.336	0.028	9	25th
rs10008582	4	146026232	Intergenic	-0.74	0.336	0.028	9	26th
rs7041171	9	111939350	Intergenic	-0.74	0.336	0.028	9	27th
rs9995307	4	146488500	*SLC10A7*	-0.74	0.336	0.028	9	28th
rs4958888	5	151093281	*TNIP1*	-0.741	0.337	0.028	9	29th
rs12991146	2	59884087	*ENSG00000271955*	-0.742	0.337	0.028	9	32nd
rs4945276	11	78453938	*NARS2*	-0.742	0.336	0.028	9	31st
rs56024498	16	76893238	*ENSG00000259995*	-0.742	0.336	0.028	9	30th
rs144476584	9	23029711	*ENSG00000284418*	-0.743	0.336	0.027	8	33rd
rs642811	11	78053929	*NDUFC2-KCTD14*	-0.743	0.336	0.027	8	34th
rs35346557	3	190120229	*P3H2*	-0.744	0.336	0.027	8	35th
rs71472777	11	24121389	Intergenic	-0.744	0.336	0.027	8	36th

rs118072482	8	138017250	*ENSG00000253288*	-0.745	0.336	0.027	8	37th
rs13410191	2	137643025	*THSD7B*	-0.745	0.336	0.027	8	38th
rs35714695	17	28392769	*SARM1*	-0.745	0.336	0.027	8	39th
rs11718653	3	122759011	*HSPBAP1*	-0.746	0.336	0.027	8	40th
rs2240601	17	57673751	*MSI2*	-0.746	0.336	0.027	8	41st
rs1570281	6	146517546	Intergenic	-0.747	0.336	0.027	8	43rd
rs74654358	12	64488187	*TBK1*	-0.747	0.337	0.027	8	42nd
rs2253050	16	74040378	*PSMD7-DT*	-0.748	0.336	0.027	8	46th
rs34432311	2	177346958	*NFE2L2*	-0.748	0.336	0.026	8	47th
rs76805704	12	64138597	*SRGAP1*	-0.748	0.337	0.027	8	45th
rs77058105	20	17070347	Intergenic	-0.748	0.336	0.027	8	44th
rs10463311	5	151031274	*TNIP1*	-0.749	0.336	0.026	8	48th
rs16905848	11	20223160	Intergenic	-0.749	0.336	0.026	7	50th
rs2477866	1	233152025	*PCNX2*	-0.749	0.337	0.026	7	49th
rs12138742	1	119591406	Intergenic	-0.75	0.336	0.026	7	52nd
rs138116283	4	4318540	*ZBTB49*	-0.75	0.336	0.026	7	53rd
rs6020200	20	50017227	Intergenic	-0.75	0.336	0.026	7	51st
rs11695294	2	176608890	*LINC01117*	-0.751	0.336	0.026	7	56th
rs60318796	17	32991387	*SPACA3*	-0.751	0.336	0.026	7	54th
rs73103977	12	53513226	*ATF7*	-0.751	0.336	0.026	7	55th
rs10430614	10	131935136	*PPP2R2D*	-0.752	0.336	0.026	7	58th
rs117452182	13	27293529	Intergenic	-0.752	0.336	0.026	7	57th
rs9956309	18	38136697	Intergenic	-0.752	0.336	0.026	7	59th
rs1159918	4	99321852	*ADH1B*	-0.753	0.336	0.025	7	60th
rs143747467	16	11367549	*RMI2*	-0.753	0.337	0.026	7	61st
rs2893656	7	106534655	*ENSG00000243797*	-0.754	0.336	0.025	7	62nd
rs2176039	22	45189151	Intergenic	-0.755	0.336	0.025	6	65th
rs28660489	12	64560784	*RASSF3*	-0.755	0.336	0.025	7	63rd
rs7602576	2	112942040	Intergenic	-0.755	0.336	0.025	7	64th
rs118038177	11	121459793	*SORL1*	-0.756	0.336	0.025	6	68th
rs6947666	7	137708989	*DGKI*	-0.756	0.336	0.025	6	66th
rs9567838	13	47359837	Intergenic	-0.756	0.336	0.025	6	67th
rs112288580	11	98729819	Intergenic	-0.757	0.336	0.025	6	70th
rs115980385	7	141463227	*TMEM178B*	-0.757	0.336	0.025	6	69th
rs6997565	8	2560505	Intergenic	-0.757	0.337	0.025	6	71st
rs72973932	11	74400615	*ENSG00000254631*	-0.758	0.337	0.025	6	72nd
rs11171999	12	56846925	*ENSG00000258679*	-0.759	0.336	0.024	6	73rd
rs141730255	7	138253804	Intergenic	-0.761	0.336	0.024	6	74th
rs111704832	15	93374070	*ENSG00000257060*	-0.762	0.336	0.024	6	77th
rs141347161	7	42377714	Intergenic	-0.762	0.336	0.024	6	76th
rs538622	5	172920676	*ERGIC1*	-0.762	0.336	0.024	6	75th
rs72792226	10	48204534	*FRMPD2*	-0.762	0.336	0.024	6	78th
rs10492593	13	66919985	*PCDH9*	-0.763	0.336	0.024	5	83rd

rs12900374	15	82741261	Intergenic	-0.763	0.336	0.024	6	80th
rs17326496	5	113340882	*MCC*	-0.763	0.336	0.023	5	81st
rs28407220	2	33816106	*LINC01320*	-0.763	0.336	0.023	5	86th
rs35059420	5	169995487	*DOCK2*	-0.763	0.336	0.023	5	84th
rs36037136	1	164710739	*PBX1*	-0.763	0.336	0.023	5	85th
rs4273590	5	159335610	*ENSG00000249738*	-0.763	0.336	0.024	5	82nd
rs9653747	21	18669100	*MIR548XHG*	-0.763	0.336	0.024	6	79th
rs12967284	18	12532099	*SPIRE1*	-0.764	0.336	0.023	5	89th
rs144049425	2	162505368	*KCNH7*	-0.764	0.336	0.023	5	87th
rs68072647	17	9224590	*NTN1*	-0.764	0.336	0.023	5	88th
rs112820958	5	79465868	*HOMER1*	-0.765	0.336	0.023	5	91st
rs12079484	1	181048307	*MR1*	-0.765	0.336	0.023	5	90th
rs144129573	12	114485139	Intergenic	-0.765	0.336	0.023	5	93rd
rs35851984	17	28239061	Intergenic	-0.765	0.336	0.023	5	92nd
rs11652752	17	67379776	*PITPNC1*	-0.766	0.336	0.023	5	94th
rs12220832	10	80806689	Intergenic	-0.766	0.336	0.023	4	99th
rs150278778	1	209379025	*LINC01698*	-0.766	0.336	0.023	4	100th
rs5766195	22	44921429	*PHF21B*	-0.766	0.336	0.023	5	95th
rs77238283	17	11797907	*DNAH9*	-0.766	0.336	0.023	4	97th
rs7764458	6	83116819	*DOP1A*	-0.766	0.336	0.023	4	101st
rs79612353	20	59884601	*SYCP2*	-0.766	0.336	0.023	5	96th
rs8101883	19	56681170	*ZIM2-AS1*	-0.766	0.336	0.023	4	98th
rs12229321	12	64518074	*RASSF3*	-0.767	0.336	0.023	4	103rd
rs1442671	18	69528820	*DOK6*	-0.767	0.336	0.023	4	102nd
rs72838433	2	127913724	Intergenic	-0.768	0.336	0.023	4	104th
rs430979	4	2812971	*SH3BP2*	-0.769	0.336	0.022	4	106th
rs6037557	20	405220	Intergenic	-0.769	0.336	0.022	4	107th
rs79446108	7	137863942	Intergenic	-0.769	0.336	0.022	4	105th
rs116488199	10	1675116	*ADARB2*	-0.77	0.336	0.022	4	108th
rs13387347	2	168898336	*SPC25*	-0.77	0.336	0.022	4	111th
rs2985994	13	45539849	*ERICH6B*	-0.77	0.336	0.022	4	112th
rs76427181	6	86684739	Intergenic	-0.77	0.336	0.022	4	109th
rs78549703	19	17638733	*UNC13A*	-0.77	0.336	0.022	4	110th
rs193044924	17	15767328	Intergenic	-0.771	0.336	0.022	3	114th
rs6603044	15	83015059	*ENSG00000259805*	-0.771	0.336	0.022	3	113th
rs10139154	14	30678292	*SCFD1*	-0.772	0.336	0.022	3	116th
rs150949995	5	150518937	*NDST1*	-0.772	0.336	0.022	3	117th
rs8053191	16	81117558	*PKD1L2*	-0.772	0.337	0.022	3	115th
rs72733862	5	8446181	*LINC02226*	-0.773	0.336	0.022	3	118th
rs76323495	16	16859460	Intergenic	-0.773	0.336	0.022	3	119th
rs4974650	4	2309992	*ZFYVE28*	-0.774	0.336	0.022	3	120th
rs79676202	12	49786775	Intergenic	-0.774	0.336	0.022	3	121st
rs16865645	2	177696567	*PDE11A*	-0.775	0.336	0.021	3	122nd

rs116946806	7	131997812	Intergenic	-0.776	0.336	0.021	3	123rd
rs320019	1	48610454	*AGBL4*	-0.776	0.336	0.021	3	124th
rs11185388	1	104198732	*ENSG00000215869*	-0.777	0.336	0.021	3	125th
rs3109207	4	168675360	*PALLD*	-0.777	0.336	0.021	3	126th
rs11065961	12	111585263	*ATXN2*	-0.778	0.336	0.021	3	127th
rs111970477	3	47064091	*SETD2*	-0.779	0.336	0.021	3	128th
rs115348904	4	158390602	*RXFP1*	-0.779	0.337	0.021	2	129th
rs112913348	5	108554187	Intergenic	-0.78	0.336	0.021	2	131st
rs72716562	5	7957484	Intergenic	-0.78	0.336	0.021	2	130th
rs8180839	7	5200339	*WIPI2*	-0.782	0.336	0.02	2	132nd
rs12472309	2	7266695	Intergenic	-0.783	0.336	0.02	2	134th
rs2271689	10	17046273	*CUBN*	-0.783	0.336	0.02	2	133rd
rs12369156	12	120729872	*ACADS*	-0.785	0.336	0.02	2	135th
rs75087725	21	44333234	*CFAP410*	-0.785	0.336	0.02	2	136th
rs10067826	5	10282407	*CMBL*	-0.786	0.336	0.02	2	138th
rs35318094	5	180245984	*MAPK9*	-0.786	0.336	0.02	2	139th
rs73152707	3	86619077	Intergenic	-0.786	0.336	0.02	2	137th
rs970258	2	5138399	Intergenic	-0.786	0.336	0.02	2	140th
rs117860708	11	1537217	Intergenic	-0.788	0.336	0.019	2	141st
rs4676496	3	39456514	Intergenic	-0.788	0.336	0.019	2	142nd
rs11067262	12	114724621	*TBX3-AS1*	-0.789	0.336	0.019	2	144th
rs3098553	15	27631110	Intergenic	-0.789	0.336	0.019	2	143rd
rs9819308	3	1691812	Intergenic	-0.79	0.336	0.019	1	145th
rs10050775	5	38007940	Intergenic	-0.791	0.336	0.019	1	146th
rs10443173	1	86068071	*COL24A1*	-0.796	0.336	0.018	1	148th
rs144387708	12	119264395	Intergenic	-0.796	0.336	0.018	1	150th
rs3798105	5	133194937	*ENSG00000248245*	-0.796	0.336	0.018	1	149th
rs9903355	17	36580791	*GGNBP2*	-0.796	0.336	0.018	1	147th
rs7209200	17	5066645	Intergenic	-0.797	0.336	0.018	1	151st
rs651001	6	11569169	*TMEM170B*	-0.798	0.336	0.018	1	153rd
rs79496463	8	132904843	*TG*	-0.798	0.336	0.018	1	152nd
rs9813285	3	29329178	*RBMS3*	-0.804	0.336	0.017	1	154th
rs118071175	14	28515686	Intergenic	-0.805	0.336	0.017	1	155th
rs7930973	11	44159478	*EXT2*	-0.806	0.336	0.017	1	156th
rs11702120	21	23786163	Intergenic	-0.808	0.336	0.016	1	157th
rs7258235	19	2612120	*GNG7*	-0.811	0.336	0.016	1	158th
rs4292737	8	10401605	*MSRA*	-0.814	0.336	0.016	1	159th
rs112348322	4	118882809	*SYNPO2*	-0.833	0.336	0.013	1	160th
rs12886280	14	31829453	*NUBPL*	-0.834	0.336	0.013	1	161st

Table S4. Enrichment analysis identified pathways influencing the age of symptom onset among *C9orf72* carriers, related to Figure 4.

Enrichment analysis for decile ten was based on the variants that make up decile ten plus the *C9orf72* gene. g:Profiler maps intronic variants to their corresponding gene. Decile refers to the decile that yielded the pathways on enrichment analysis. No significant pathways were identified for deciles 1–4 and 9. GO:MF, Gene Ontology, molecular function; GO:BP, Gene Ontology, biological process; KEGG, Kyoto Encyclopedia of Genes and Genomes.

Source	Term name	Category	Term ID	Adjusted p-value	Decile
GO:MF	ATP-dependent microtubule motor activity	Transport	GO:1990939	0.020	10
GO:MF	Motor activity	Transport	GO:0003774	0.020	10
GO:MF	Microtubule motor activity	Transport	GO:0003777	0.020	10
GO:BP	Vacuolar transport	Transport	GO:0007034	0.043	10
GO:BP	Cytosolic transport	Transport	GO:0016482	0.045	10
GO:MF	Tubulin binding	Transport	GO:0015631	0.020	10
GO:BP	Developmental growth involved in morphogenesis	Transport	GO:0060560	0.034	10
GO:BP	Developmental cell growth	Transport	GO:0048588	0.034	10
GO:BP	Cell growth	Transport	GO:0016049	0.040	10
GO:BP	Synaptic vesicle transport	Axonal transport	GO:0048489	0.034	10
GO:BP	Synaptic vesicle localization	Axonal transport	GO:0097479	0.035	10
GO:BP	Axon extension	Axonal transport	GO:0048675	0.040	10
GO:BP	Neuron projection extension	Axonal transport	GO:1990138	0.028	10
GO:BP	Axonogenesis	Axonal transport	GO:0007409	0.040	10
GO:BP	Lysosomal transport	Autophagy	GO:0007041	0.040	10
GO:BP	Positive regulation of autophagy	Autophagy	GO:0010508	0.040	10
GO:BP	Establishment of protein localization to mitochondrion	Mitochondria	GO:0072655	0.040	10
GO:BP	Protein targeting to mitochondrion	Mitochondria	GO:0006626	0.040	10
GO:BP	Protein localization to mitochondrion	Mitochondria	GO:0070585	0.041	10
KEGG	Amyotrophic lateral sclerosis	Neurodegeneration	KEGG:05014	0.001	10
KEGG	Pathways of neurodegeneration - multiple diseases	Neurodegeneration	KEGG:05022	0.002	10
GO:MF	Oxidoreductase activity, acting on the CH-OH group of donors, NAD or NADP as acceptor	Other	GO:0016616	0.020	10
GO:MF	Oxidoreductase activity, acting on CH-OH group of donors	Other	GO:0016614	0.021	10
GO:BP	Phenol-containing compound metabolic process	Other	GO:0018958	0.040	10
GO:BP	Regulation of innate immune response	Defense response	GO:0045088	0.011	8
GO:BP	Positive regulation of response to biotic stimulus	Response to biotic stimulus	GO:0002833	0.017	8
GO:BP	Innate immune response-activating signaling pathway	Defense response	GO:0002758	0.017	8
GO:BP	Positive regulation of innate immune response	Defense response	GO:0045089	0.017	8
GO:BP	Pattern recognition receptor signaling pathway	Defense response	GO:0002221	0.017	8

GO:BP	Activation of innate immune response	Defense response	GO:0002218	0.017	8
GO:BP	Toll-like receptor signaling pathway	Receptor signaling pathway	GO:0002224	0.017	8
GO:BP	Immune response-activating signaling pathway	Signal transduction	GO:0002757	0.023	8
GO:BP	Positive regulation of defense response	Defense response	GO:0031349	0.023	8
GO:BP	Immune response-regulating signaling pathway	Signal transduction	GO:0002764	0.025	8
GO:MF	2-oxoglutarate-dependent dioxygenase activity	Oxidoreductase activity	GO:0016706	0.031	8
GO:MF	Dioxygenase activity	Oxidoreductase activity	GO:0051213	0.037	8
GO:BP	Stem cell differentiation	Cellular developmental process	GO:0048863	0.049	8
GO:MF	Transcription coactivator binding	Transcription factor binding	GO:0001223	0.006	7
GO:MF	Transcription coregulator binding	Transcription factor binding	GO:0001221	0.018	7
GO:BP	Neutral lipid metabolic process	Cellular lipid metabolic process	GO:0006638	0.016	6
GO:BP	Acylglycerol metabolic process	Cellular lipid metabolic process	GO:0006639	0.016	6
GO:BP	Endoplasmic reticulum to Golgi vesicle-mediated transport	Intracellular transport	GO:0006888	0.016	6
GO:MF	Small GTPase binding	Enzyme binding	GO:0031267	0.019	6
GO:MF	GTPase binding	Enzyme binding	GO:0051020	0.019	6
GO:MF	Molecular function inhibitor activity	Molecular function regulatory activity	GO:0140678	0.019	6
GO:MF	Enzyme inhibitor activity	Enzyme regulator activity	GO:0004857	0.019	6
GO:BP	Golgi vesicle transport	Vesicle-mediated transport	GO:0048193	0.026	6
GO:BP	Glycerolipid metabolic process	Cellular lipid metabolic process	GO:0046486	0.030	6
GO:MF	T cell receptor binding	Signaling receptor binding	GO:0042608	0.002	5

Table S5. The list of genes used in the gene-gene similarity network *GREP* analysis, related to Figure 5.

Seed genes refer to the genetically identified genes in decile ten.

Rank	Gene
Seed gene	*ABLIM2*
Seed gene	*C9orf72*
Seed gene	*DIP2B*
Seed gene	*DNAH17*
Seed gene	*KIF5A*
Seed gene	*LINC00598*
Seed gene	*ME2*
Seed gene	*OCA2*
Seed gene	*PRKN*
Seed gene	*SAMM50*
Seed gene	*SDR9C7*
1	*RIMS3*
2	*CAMKV*
3	*CLASP2*
4	*TUBB4A*
5	*KCNQ2*
6	*BSN*
7	*KCNC1*
8	*MAPK4*
9	*SNCB*
10	*CLVS2*
11	*CNTN2*
12	*SEPTIN3*
13	*NCAN*
14	*KIF1A*
15	*LGI3*
16	*CDH22*
17	*CDHR1*
18	*ADCYAP1R1*
19	*OTUD7A*
20	*KIF1B*
21	*SLC8A2*
22	*TMEM151B*
23	*TRIM9*
24	*PHF24*
25	*VSTM2B*
26	*CACNA1B*
27	*DOCK3*

28	*NDRG4*
29	*ATP1A3*
30	*GPM6B*
31	*CRHR1*
32	*RBFOX1*
33	*TPPP*
34	*GRIA2*
35	*NEFL*
36	*SLC35F1*
37	*SNAP91*
38	*KCNJ9*
39	*HMP19*
40	*ACTL6B*
41	*CARMIL3*
42	*PLP1*
43	*SYP*
44	*CELSR2*
45	*SULT4A1*
46	*NEFM*
47	*RIMBP2*
48	*CPLX2*
49	*ASTN1*
50	*CNTN1*
51	*SHISA7*
52	*STMN4*
53	*SYN2*
54	*GNAL*
55	*DLG2*
56	*KCNA2*
57	*ARHGEF4*
58	*KIF5C*
59	*IGSF11*
60	*NSG1*
61	*CELF3*
62	*ZDHHC22*
63	*HRH3*
64	*GDAP1L1*
65	*SCRT1*
66	*INA*
67	*NRXN1*
68	*WNK2*
69	*PTPN5*
70	*MMD2*

71	*HPCAL4*
72	*CHRNB2*
73	*JPH3*
74	*ATP2B3*
75	*MAPK8IP2*
76	*HAPLN2*
77	*MYT1*
78	*PAFAH1B1*
79	*CDK5R1*
80	*NOL4*
81	*SYNPR*
82	*ZNF536*
83	*GALNT8*
84	*CTNNA2*
85	*ST8SIA3*
86	*ABCG4*
87	*LINGO1*
88	*ADAM22*
89	*AMER2*
90	*PAK5*
91	*PPFIA3*
92	*SH3GL2*
93	*PREPL*
94	*ELMOD1*
95	*ATP2B2*
96	*LRRC4B*
97	*SORCS1*
98	*DUSP26*
99	*JPH4*
100	*MAST1*
101	*BRSK2*
102	*CA11*
103	*ELAVL3*
104	*GNG3*
105	*STXBP5L*
106	*PTPRZ1*
107	*MAPT*
108	*ZCCHC12*
109	*RIMS4*
110	*SYT4*
111	*TMEM179*
112	*CLASP1*
113	*ELAVL4*

114	*SOCS7*
115	*ADGRB3*
116	*RUNDC3A*
117	*SCN3B*
118	*PSD2*
119	*EFR3B*
120	*GARNL3*
121	*GRID1*
122	*ZDHHC11B*
123	*TMEM63C*
124	*RAB3C*
125	*CHD5*
126	*ADGRL3*
127	*DPYSL5*
128	*GRIK3*
129	*PHYHIPL*
130	*CCDC177*
131	*GRIA4*
132	*FBXL16*
133	*IGSF21*
134	*SEZ6L*
135	*TTBK1*
136	*CNTFR*
137	*NRXN2*
138	*LANCL1*
139	*OLFM3*
140	*SCG3*
141	*ABCC8*
142	*UNC13A*
143	*SH3GL3*
144	*GRIK5*
145	*NELL1*
146	*ATP1B2*
147	*SCN4B*
148	*CADM2*
149	*RIPPLY2*
150	*B3GAT1*
151	*IGSF9B*
152	*ANK2*
153	*CADM4*
154	*UBE2QL1*
155	*PCDH8*
156	*SYN1*

157	*ADGRA1*
158	*ADAM11*
159	*KCNB2*
160	*SMIM10L2A*
161	*NXPH1*
162	*CHRNA4*
163	*NAPB*
164	*GAP43*
165	*MAP3K9*
166	*PGBD5*
167	*LRRTM3*
168	*PIP4K2B*
169	*RAB3A*
170	*SLIT1*
171	*TCEAL5*
172	*GABRG2*
173	*GNAO1*
174	*GRID2*
175	*PPP2R2C*
176	*RUFY3*
177	*CASKIN1*
178	*ADGRL1*
179	*SORCS3*
180	*SPTBN4*
181	*CNTNAP4*
182	*PDZD4*
183	*PEX5L*
184	*SOX8*
185	*CSPG5*
186	*ATCAY*
187	*RPRD1A*
188	*RGS8*
189	*PHF21B*
190	*ACTN2*
191	*GAD2*
192	*SLC6A11*
193	*SLC32A1*
194	*KCNJ4*
195	*MEGF11*
196	*POU3F3*
197	*SV2B*
198	*ELAVL2*
199	*IGLON5*

200	*LRRC3B*

Table S6. The list of repurposable, approved drugs identified by the GREP analysis, related to Figure 5.

Drug indications and mechanisms of action were curated from the Drugbank database [S15]. GABA, gamma-aminobutyric acid; GABA(A), gamma-aminobutyric acid Type A; VDCC, Voltage-dependent calcium channel; KCNK3, Potassium Two Pore Domain Channel Subfamily K Member 3; CHRNA4, Cholinergic Receptor Nicotinic Alpha 4 Subunit; CaV2.2, Neuronal voltage-gated N-type Calcium Channel.

Drug	Indication	Mechanism of action
Acamprosate	Withdrawal symptoms of alcoholism	Analogue of GABA
Adinazolam	Seizures	GABA positive allosteric modulator
Alprazolam	Anxiety and panic disorders	GABA(A) receptor positive allosteric modulator
Betahistine	Vertigo	H1-receptor agonist
Brivaracetam	Seizures	Unknown, synaptic GABA release
Bromazepam	Anxiety and panic disorders	GABA(A) receptor positive allosteric modulator
Chlordiazepoxide	Withdrawal symptoms of alcoholism	GABA(A) receptor positive allosteric modulator
Cinolazepam	Sleep disorders	GABA positive allosteric modulator
Clobazam	Seizures	GABA positive allosteric modulator
Clonazepam	Anxiety and panic disorders	GABA(A) receptor positive allosteric modulator
Clotiazepam	Anxiety	GABA(A) receptor positive allosteric modulator
Desflurane	Anesthetic	GABA(A) receptor positive allosteric modulator
Diazepam	Anxiety and alcohol withdrawal	GABA(A) receptor positive allosteric modulator
Enflurane	Anesthetic	GABA(A) receptor potentiator
Estazolam	Insomnia	GABA(A) receptor positive allosteric modulator
Eszopiclone	Insomnia	GABA(A) receptor potentiator
Ethchlorvynol	Insomnia	GABA(A) receptor positive allosteric modulator
Etizolam	Anxiety and insomnia	GABA(A) receptor positive allosteric modulator
Etomidate	Anesthetic	GABA receptor subunit alpha-1 potentiator
Fludiazepam	Convulsion	GABA receptor subunit alpha-1 agonist
Flurazepam	Anxiety and convulsion	GABA(A) receptor positive allosteric modulator
Gabapentin	Convulsion	VDCC subunit alpha-2/delta-1 inhibitor
Glutethimide	Sedative	GABA receptor subunit alpha-1 agonist
Halazepam	Seizures and anxiety	GABA(A) receptor positive allosteric modulator
Halothane	Anesthetic	KCNK3 binder
Isoflurane	Anesthetic	GABA receptor subunit alpha-1 agonist
Levetiracetam	Seizures	CaV2.2 subunit alpha-1B inhibitor
Lorazepam	Seizures, anxiety, and panic disorders	GABA(A) receptor positive allosteric modulator
Lormetazepam	Anxiety	GABA(A) receptor positive allosteric modulator
Meprobamate	Anxiety	GABA agonist
Metharbital	Convulsion	GABA receptor subunit alpha-2 potentiator
Methoxyflurane	Anesthetic	GABA receptor subunit alpha-1 agonist
Methylphenobarbital	Seizures	Depressant of the central nervous system
Midazolam	Anxiety and convulsion	GABA(A) receptor positive allosteric modulator
Nicotine	Smoking cessation.	Neuronal CHRNA4 agonist

Nitrazepam	Anxiety and insomnia	GABA positive allosteric modulator
Oxazepam	Withdrawal symptoms of alcoholism	GABA(A) receptor positive allosteric modulator
Pentobarbital	Seizures and sedation	GABA(A) receptor potentiator
Pitolisant	Narcolepsy	Antagonist at the histamine H3 receptor
Prazepam	Anxiety	GABA(A) receptor positive allosteric modulator
Primidone	Seizures	GABA receptor subunit beta-2 potentiator
Propofol	Sedative	GABA receptor subunit beta-2 potentiator
Quazepam	Insomnia	GABA positive allosteric modulator
Sevoflurane	Anesthetic	GABA(A) receptor agonist
Stiripentol	Seizures	GABA(A) receptor agonist allosteric modulator
Talbutal	Sedative	GABA receptor subunit alpha-2 potentiator
Temazepam	Anxiety and panic disorders	GABA(A) receptor positive allosteric modulator
Topiramate	Seizures	GABRA1 agonist
Triazolam	Insomnia	GABA(A) receptor positive allosteric modulator
Varenicline	Smoking cessation	CHRNA4 partial agonist
Ziconotide	Chronic pain	CaV2.2 subunit alpha-1B inhibitor

Table S7. *KIF5A* genotypes influenced the age of symptom onset among *C9orf72* carriers, related to Figure 4.

The table summarizes the age at onset of *C9orf72* individuals carrying the rs113247976 (chr12:57581917) variant in the *KIF5A* gene. N, number of cases.

Variant	Status	Genotype	Age at onset mean	Age at onset standard deviation	Cases (n)
rs113247976 (*KIF5A*)	*C9orf72* carrier	CC	57.81	9.49	780
		CT	54.19	11.09	37
		TT	NA	NA	0
	Non-carrier	CC	60.08	12.54	6,827
		CT	60.06	12.14	206
		TT	61.75	11.76	4

Table S8. The sources of the cohorts used in this study, related to Star Methods.

Dataset	Sample size	Reference
Reference	12,577 ALS cases & 23,475 controls	16
Training	7,030 ALS cases & 34,235 controls	16,17
Test (*C9orf72*)	817 ALS/FTD	16,18
Replication (*C9orf72*)	699 ALS/FTD	14

Table S9. The clinical descriptions of the test cohort, related to Star Methods.

The test cohort is composed of 817 *C9orf72* carriers. FTLD, frontotemporal lobar degeneration; FTLD-Tau, frontotemporal lobar degeneration with tau pathology; PSP, progressive supranuclear palsy; FTLD TDP, frontotemporal lobar degeneration with TDP-43 inclusions; FTLD-U, frontotemporal lobar degeneration with ubiquitin-positive inclusions; NOS, not otherwise specified. *One FTLD NOS patient was initially clinically misdiagnosed as PSP.

Type	Subtype	Size
Asymptomatic		13
Clinical	ALS	666
Clinical	FTD Motor neuron disease	81
Clinical	FTD Behavioral variant	23
Clinical	FTD Language variant NOS	1
Clinical	FTD Nonfluent variant	3
Clinical	FTD NOS	2
Clinical	FTD Semantic variant	3
Pathological	FTLD NOS*	9
Pathological	FTLD Tau (PSP)	2
Pathological	FTLD TDP Type A	4
Pathological	FTLD TDP Type B	1
Pathological	FTLD TDP Type Unknown	6
Pathological	FTLD U	3

Table S10. The cell lines used for the drug validation experiments, related to Figure 6 and Figure S7.

F, female; M, male.

iPSC cell line	Source	Clinical remarks	Mutation	Race	Sex	Age at collection	Supplier
CS14iCTR-nxx	Fibroblast	Clinically normal and healthy volunteer	Unknown	Caucasian	F	52	Cedars-Sinai
GM23338	Fibroblast	Clinically normal and healthy volunteer	Unknown	Caucasian	M	55	Coriell Biorepository
MIFF1	Fibroblast	Clinically normal and healthy volunteer	Unknown	Caucasian	M	<1 month old	University of Sheffield
CS02iCTR-NTn1	PBMC	Clinically normal and healthy volunteer	Unknown	Caucasian	M	51	Cedars-Sinai
ALS-183-C9	Fibroblast	ALS, age of onset = 48; disease duration = 27 months.	*C9orf72* repeat expansion	Caucasian	M	50	University of Sheffield
ALS-78	Fibroblast	ALS, age of onset = unknown; disease duration = 31.7 months	*C9orf72* repeat expansion	Caucasian	M	66	University of Sheffield
CS28iALS-C9nxx	Fibroblast	ALS, age of onset = 46; site of onset = left upper extremity	*C9orf72* repeat expansion	Caucasian	M	47	Cedars-Sinai
CS29iALS-C9nxx	Fibroblast	ALS, age of onset = unknown; disease duration = unknown	*C9orf72* repeat expansion	Caucasian	M	47	Cedars-Sinai
CS52iALS-C9nxx	Fibroblast	ALS, age of onset = 57; disease duration = 48 months	*C9orf72* repeat expansion	Unknown	M	49	Cedars-Sinai
CS29iALS-C9n1.ISOxx	Fibroblast	ALS, Age of onset: Unknown; disease duration, Unknown	Isogenic control line of CS29iALS-C9nxx	Caucasian	M	47	Cedars-Sinai
CS52iALS-C9n6.ISOxx	Fibroblast	ALS, age of onset = 57; disease duration = 48 months	Isogenic control line of CS52iALS-nxx	Unknown	M	49	Cedars-Sinai

Table S11. The antibodies used in motor neuron staining, related to Star Methods.

ChAT, Choline acetyltransferase; MAP2, Microtubule-associated protein 2; Islet 1/2, ISL LIM Homeobox 1/2.

	Host	Dilution	Wavelength (nm)	Supplier	Catalogue Number
Primary Antibodies					
Beta III tubulin	Mouse	1:1000	-	Biolegend	801201
Caspase-3	Rabbit	1:200	-	Merck Millipore	AB3623
ChAT	Goat	1:100	-	Merck Millipore	AB144P
MAP-2	Guinea pig	1:1000	-	Synaptic systems	188004
Islet 1/2	Rabbit	1:500	-	Abcam	ab109517
NeuN	Mouse	1:1000	-	Millipore	MAB377
Secondary Antibodies					
Anti-rabbit	Donkey	-	488	Thermofisher	A21206
Anti-rabbit	Donkey	-	568	Thermofisher	A10042
Anti-mouse	Donkey	-	568	Thermofisher	A10037
Anti-mouse	Donkey	-	488	Thermofisher	A21202
Anti-goat	Donkey	-	555	Thermofisher	A21432
Anti-guinea pig	Goat	-	647	Thermofisher	A21450

OTHER INFORMATION

Consortia

The members of the International ALS Genomics Consortium are:

Robert H. Baloh[1], Robert Bowser[2], Christopher B. Brady[3], Alexis Brice[4,5], James Broach[6], William Camu[7], Ruth Chia[8], Adriano Chiò[9,10,11], John Cooper-Knock[12], Daniele Cusi[13], Jinhui Ding[14], Carsten Drepper[15], Vivian E. Drory[16], Travis L. Dunckley[17], Eva Feldman[18], Mary Kay Floeter[19], Pietro Fratta[20], Glenn Gerhard[17], J. Raphael Gibbs[14], Summer B. Gibson[21], Jonathan D. Glass[22], Stephen A. Goutman[18], John Hardy[23], Matthew B. Harms[24], Terry D. Heiman-Patterson[25,26], Lilja Jansson[27], Janine Kirby[28], Hannu Laaksovirta[27], John E. Landers[29], Francesco Landi[30], Isabelle Le Ber[31], Serge Lumbroso[32], Claire Guissart[32], Daniel JL. MacGowan[33], Nicholas J. Maragakis[34], Gabriele Mora[9], Kevin Mouzat[32], Liisa Myllykangas[27], Richard W. Orrell[35], Lyle W. Ostrow[26], Stuart Pickering-Brown[36], Erik P. Pioro[37], Stefan M. Pulst[21], John M. Ravits[38], Alan E. Renton[39], Wim Robberecht[40], Ekaterina Rogaeva[41], Jeffrey D. Rothstein[42], Erika Salvi[43], Sonja W. Scholz[42,44], Michael Sendtner[45], Pamela J. Shaw[12], Katie C. Sidle[23], Zachary Simmons[46], David J. Stone[47], Pentti J. Tienari[27], Bryan J. Traynor[8,23,42,48,49], John Q. Trojanowski[50], Juan C. Troncoso[51], Miko Valori[27], Philip Van Damme[41,52], Vivianna M. Van Deerlin[50], Ludo Van Den Bosch[41], Lorne Zinman[53]

1. Department of Neurology, Cedars-Sinai Medical Center, Los Angeles, CA 90048, USA
2. Division of Neurology, Barrow Neurological Institute, Phoenix, AZ 85013, USA
3. Research and Development Service, Veterans Affairs Boston Healthcare System, Boston, MA 02130, USA
4. Centre de Recherche de l'Institut du Cerveau et de la Moelle épinière, Université Pierre et Marie Curie, Paris, France
5. INSERM U975, Paris, France
6. Department of Biochemistry, Penn State College of Medicine, Hershey, PA 17033, USA
7. ALS reference center, Gui de Chauliac Hospital, CHU and Univ Montpellier, Montpellier, France
8. Neuromuscular Diseases Research Section, National Institute on Aging, Bethesda, MD 20892, USA
9. 'Rita Levi Montalcini' Department of Neuroscience, University of Turin, Via Verdi 8, Turin, 10124, Italy
10. Neuroscience Institute of Torino, University of Turin, Turin, 10124, Italy
11. Institute of Cognitive Sciences and Technologies, C.N.R., Rome, Italy
12. Sheffield Institute for Translational Neuroscience (SITraN), University of Sheffield, Sheffield, UK
13. Bio4Dreams Scientific Unit Bio4Dreams - Business Nursery for Life Sciences Milano Italy
14. Computational Biology Group, Laboratory of Neurogenetics, National Institute on Aging, Bethesda, MD 20892, USA
15. Institute of Clinical Neurobiology, University Hospital Wuerzburg, Wuerzburg 97080, Germany
16. Department of Neurology, Tel-Aviv Sourasky Medical Center, Tel-Aviv,
17. Department of Pathology, Penn State College of Medicine, Hershey, PA 17033, USA
18. Department of Neurology, University of Michigan, 1500 E Medical Center Dr, Ann Arbor, MI 48109, USA
19. Motor Neuron Disorders Unit, National Institute of Neurological Disorders and Stroke, Bethesda, MD 20892, USA
20. Sobell Department of Motor Neuroscience and Movement Disorders, Institute of Neurology, University College London, London, WC1N 3BG, UK
21. Department of Neurology, University of Utah School of Medicine, 175 North Medical Drive East, Salt Lake City, UT 84132, USA
22. Department of Neurology, Emory University School of Medicine, Atlanta, GA 30322, USA

23. Department of Molecular Neuroscience and Reta Lila Weston Laboratories, Institute of Neurology, University College London, London, WC1N 3BG, UK
24. Department of Neurology, Columbia University, New York, NY 10032, USA
25. Department of Neurology, Drexel University College of Medicine, Philadelphia, PA 19102, USA
26. Department of Neurology, Temple University, 7602 Central Ave, Philadelphia, PA 19111, USA
27. Department of Neurology and HUSLAB, Helsinki University Hospital, Translational Immunology, Research Programs Unit and Department of Pathology University of Helsinki, Helsinki, FIN-02900, Finland
28. Department of Neuroscience, University of Sheffield, Sheffield, S10 2HQ, UK
29. Department of Neurology, University of Massachusetts Medical School, Worcester, MA 01605, USA
30. Fondazione Policlinico Universitario Agostino Gemelli IRCCS, Università Cattolica del Sacro Cuore, Rome, Italy
31. Service de Biochimie, CHU de Nîmes, Nîmes, France
32. CHU Nimes, Univ. Montpellier, INM, INSERM, Montpellier, France
33. New York Hospital Cornell University Medical Center 1305 York Avenue NYC NY 10021
34. Department of Neurology, Johns Hopkins University, Baltimore, MD 21287, USA 40.
35. Department of Clinical Neuroscience, Institute of Neurology, University College London, London, NW2 2PG, UK
36. Faculty of Human and Medical Sciences, University of Manchester, Manchester, M13 9PT, UK
37. Department of Neurology, Cleveland Clinic, Cleveland, OH 44195, USA
38. Neuroscience and Disease, University of California San Diego, 9500 Gilman Drive, La Jolla, CA 92093, USA
39. Department of Neuroscience, Ronald M. Loeb Center for Alzheimer's Disease, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA
40. Department of Neurosciences, Experimental Neurology, and Leuven Research Institute for Neuroscience and Disease, University of Leuven, Leuven, 3000, Belgium
41. Division of Neurology, Tanz Centre for Research of Neurodegenerative Diseases and Toronto Western Hospital, University of Toronto, Toronto, M5S 3H2, Canada
42. Department of Neurology, Johns Hopkins University, Baltimore, MD 21287, USA 40
43. Neurology and Headache Unit, Fondazione IRCCS Istituto Neurologico "Carlo Besta", Milan, Italy
44. Neurodegenerative Diseases Research Section, National Institute of Neurological Disorders and Stroke, Bethesda, MD 20892, USA
45. Department of Neurology, Institute for Clinical Neurobiology, University of Würzburg, Würzburg, D-97078, Germany
46. Department of Neurology, Penn State College of Medicine, Hershey, PA 17033, USA
47. Genetics, Genetics and Pharmacogenomics, Merck Research Laboratories, Merck & Co., Inc., West Point, PA 19486, USA
48. National Institute of Neurological Disorders and Stroke, Bethesda, MD 20892, USA
49. RNA Therapeutics Laboratory, National Center for Advancing Translational Sciences, NIH, Rockville, MD 20850, USA
50. Department of Pathology and Laboratory Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA
51. Clinical and Neuropathology Core, Johns Hopkins University, Baltimore, MD 21287, USA
52. VIB, Center for Brain & Disease Research, Laboratory of Neurobiology, University of Leuven, Leuven, 3000, Belgium
53. Division of Neurology, Sunnybrook Health Sciences Centre, University of Toronto, Toronto, M4N 3M5, Canada

The members of the ITALSGEN Consortium are:

Stefania M. Angelocola[1], Francesco P. Ausiello[2], Marco Barberis[3], Ilaria Bartolomei[4], Stefania Battistini[5], Enrica Bersano[6,7], Giulia Bisogni[8], Giuseppe Borghero[9], Maura Brunetti[10], Corrado Cabona[11], Andrea Calvo[10,12], Fabrizio Canale[13], Antonio Canosa[10,12,14], Teresa A. Cantisani[15], Margherita Capasso[16], Claudia Caponnetto[11], Patrizio Cardinali[1], Paola Carrera[17], Federico Casale[10], Adriano Chiò[10,12,14], Tiziana Colletti[18], Francesca L. Conforti[19], Amelia Conte[8], Elisa Conti[20,21], Massimo Corbo[22], Stefania Cuccu[9], Eleonora Dalla Bella[6], Eustachio D'Errico[23], Giovanni DeMarco[10], Raffaele Dubbioso[2], Carlo Ferrarese[20,21], Pilar M. Ferraro[11], Massimo Filippi[24,25,26,27], Nicola Fini[28], Gianluca Floris[9], Giuseppe Fuda[10], Salvatore Gallone[10], Giulia Gianferrari[28], Fabio Giannini[5], Maurizio Grassano[10], Lucia Greco[29], Barbara Iazzolino[10], Alessandro Introna[23], Vincenzo La Bella[18], Serena Lattante[30,31], Giuseppe Lauria[6,32], Rocco Liguori[33], Giancarlo Logroscino[34,35], Francesco O. Logullo[36], Christian Lunetta[29] , Paola Mandich[11,37], Jessica Mandrioli[38,39], Umberto Manera[10], Fiore Manganelli[2], Giuseppe Marangi[30,31], Kalliopi Marinou[40], Maria Giovanna Marrosu[41], Ilaria Martinelli[28], Sonia Messina[42], Cristina Moglia[10,12], Maria Rosaria Monsurrò[43], Gabriele Mora[9], Lorena Mosca[44], Maria R. Murru[9], Paola Origone[11], Carla Passaniti[43], Cristina Petrelli[36], Antonio Petrucci[45], Angelo Pirisi[46], Susanna Pozzi[29], Maura Pugliatti[46], Angelo Quattrini[47], Claudia Ricci[5], Giulia Riolo[5], Nilo Riva[47], Massimo Russo[48], Mario Sabatelli[49], Paolina Salamone[10], Marco Salivetto[29], Fabrizio Salvi[4], Marialuisa Santarelli[50], Luca Sbaiz[3], Riccardo Sideri[40], Isabella Simone[23], Cecilia Simonini[28], Rossella Spataro[18], Raffaella Tanel[51], Gioacchino Tedeschi[43], Anna Ticca[52], Antonella Torriello[53], Stefania Tranquilli[9], Lucio Tremolizzo[20,21], Francesca Trojsi[43], Rosario Vasta[10], Veria Vacchiano[4], Giuseppe Vita[48], Paolo Volanti[54], Marcella Zollino[55,56], Elisabetta Zucchi[28]

1. Neurology Unit, AST Ferme, Marche, Italy
2. Department of Neurosciences, Reproductive Sciences and Odontostomatology, University of Naples Federico II, Napoli, Italy
3. Department of Medical Genetic, Azienda Ospedaliero Universitaria Città Della Salute e Della Scienza, Torino, Italy
4. Center for Diagnosis and Cure of Rare Diseases, Department of Neurology, IRCCS Institute of Neurological Sciences, Bologna, Italy
5. Department of Medical, Surgical and Neurological Sciences, University of Siena, Siena, Italy
6. 3rd Neurology Unit and Motor Neuron Diseases Centre, IRCCS Foundation "Carlo Besta" Neurological Institute, Milano, Italy
7. 'L. Sacco' Department of Biomedical and Clinical Sciences, Università degli Studi di Milano, Università degli Studi di Milano, Milano, Italy
8. NeuroMuscular Omnicentre (NEMO), Serena Onlus, Foundation- Pol. A. Gemelli, Roma, Italy
9. Neurologic Unit, Monserrato University Hospital, Cagliari University, Cagliari, Italy
10. `Rita Levi Montalcini` Department of Neuroscience, Amyotrophic Lateral Sclerosis Center, University of Torino, Torino, Italy
11. Department of Neurosciences, Rehabilitation, Ophthalmology, Genetics and Maternal-Child Sciences, IRCCS Ospedale Policlinico San Martino, Genoa, Italy
12. Division of Neurology, Azienda Ospedaliero-Universitaria Città della Salute e della Scienza di Torino, Torino, Italy
13. Department of Advanced Medical and Surgical Sciences, University of Campania "Luigi Vanvitelli", Caserta, Italy
14. Institute of Cognitive Sciences and Technologies, C.N.R., Rome, Italy
15. Struttura complessa di Neurofisiopatologia, Azienda Ospedaliera di Perugia, Perugia, Italy
16. Unit of Neurology, Ospedale Clinicizzato SS Annunziata, Chieti, Italy
17. Unit of Genomics for the diagnosis of human pathologies, IRCCS San Raffaele Scientific Institute, Milano, Italy
18. ALS Clinical Research Center, Bi.N.D., University of Palermo, Palermo, Italy

19. Department of Pharmacy, Health and Nutritional Sciences, University of Calabria, Rende, Italy
20. Neurology Unit, "San Gerardo" hospital, Monza, Italy
21. School of Medicine and Surgery and Milan Center for Neuroscience (NeuroMI), University of Milano-Bicocca, Milano, Italy
22. Department of Neurorehabilitation Sciences, Casa Cura Policlinico, Milano, Italy
23. Department of Basic Medical Sciences, Neurosciences and Sense Organs, University of Bari "Aldo Moro", Policlinic, Bari, Italy
24. Neurology Unit and Rehabilitation Unit, IRCCS "San Raffaele Scientific Institute", Milano, Italy
25. Neuroimaging Research Unit, Division of Neuroscience, IRCCS "San Raffaele Scientific Institute", Milano, Italy
26. Neurophysiology Service, IRCCS "San Raffaele Scientific Institute," Milano, Italy
27. Vita-Salute San Raffaele University, Milano, Italy
28. Department of Neurosciences, Ospedale Civile S. Agostino Estense, Azienda Ospedaliero Universitaria di Modena, Modena, Italy
29. NEMO Clinical Center Milano, Fondazione Serena Onlus, Milano, Italy
30. Genetica Medica, Fondazione Policlinico Universitario A. Gemelli IRCCS, Roma, Italy
31. Dipartimento Universitario Scienze della Vita e Sanità Pubblica, Sezione di Medicina Genomica, Università Cattolica del Sacro Cuore Facoltà di Medicina e Chirurgia, Roma, Italy
32. Department of Biomedical and Clinical Sciences Luigi Sacco, University of Milan, Milano, Italy
33. Department of Biomedical and Neuromotor Sciences, University of Bologna, Bologna, Italy
34. Department of Basic Medical Sciences, Neuroscience and Sense Organs, University of Bari Aldo Moro, Bari, Italy
35. Center for Neurodegenerative Diseases and the Aging Brain, Department of Clinical Research in Neurology, Pia Fondazione Cardinale G Panico, Tricase, Italy
36. Neurology Unit, AV3, ASUR Marche, Macerata, Italy
37. Medical Genetics Unit, IRCCS Ospedale Policlinico San Martino, Genoa, Italy
38. Department of Neurosciences, Azienda Ospedaliero Universitaria di Modena, Modena, Italy
39. Department of Biomedical, Metabolic and Neural Sciences, University of Modena and Reggio Emilia, Modena, Italy
40. Department of Neurorehabilitation, Istituti Clinici Scientifici Maugeri IRCCS, Institute of Milan, Milano, Italy
41. Department of Neurology, Azienda Universitario Ospedaliera di Cagliari and University of Cagliari, Cagliari, Italy
42. NEuroMuscular Omnicentre (NEMO) Sud, Fondazione Aurora, OUC Neurology and Neuromuscular Disorders, University of Messina, Italy, Messina, Italy
43. Department of Advanced Medical and Surgical Sciences, University of Campania "Luigi Vanvitelli", Napoli, Italy
44. Department of Laboratory Medicine, Medical Genetics, Niguarda Ca' Granda Hospital, Milano, Italy
45. Neurology Department, San Camillo Hospital, Roma, Italy
46. Department of Biomedical and Surgical Sciences, Section of Neurological, Psychiatric and Psychological Sciences, University of Ferrara, Ferrara, Italy
47. Department of Neurology, IRCCS "San Raffaele Scientific Institute", Milano, Italy
48. OUC Neurology and Neuromuscular Disorders, University of Messina, Messina, Italy
49. NeuroMuscular Omnicentre (NEMO) - Fondazione Policlinico Universitario A. Gemelli, Università Cattolica del Sacro Cuore, Roma, Italy
50. Department of Medicine, Azienda Complesso Ospedaliero, San Filippo Neri, Roma, Italy
51. Operative Unit of Neurology, S. Chiara Hospital, Trento, Italy
52. Department of Neurology, Azienda Ospedaliera San Francesco, Nuoro, Italy
53. ALS Center, Operative Unit of Neurology, AOU "San Giovanni di Dio e Ruggi d'Aragona", Salerno, Italy
54. Neurorehabilitation Unit – ALS Center, Istituti Clinici Scientifici (ICS) Maugeri, Mistretta, Italy

55. Sezione di Medicina Genomica, Dipartimento Scienze della Vita e Sanità Pubblica, Facoltà di Medicina e Chirurgia, Università Cattolica Sacro Cuore, Roma, Italy
56. Unità di Genetica Medica, Fondazione Policlinico Universitario A. Gemelli IRCCS, Roma, Italy

The members of the SLAGEN Consortium are:

Vincenzo Silani[1,2], Isabella Fogh[3], Nicola Ticozzi[1,2], Antonia Ratti[2,4], Cinzia Tiloca[2], Silvia Peverelli[2], Cinzia Gellera[5], Giuseppe Lauria Pinter[6,7], Franco Taroni[5], Viviana Pensato[5], Barbara Castellotti[5], Giacomo P. Comi[1,8], Stefania Corti[1,8], Roberto Del Bo[1,8], Cristina Cereda[9], Mauro Ceroni[9,10], Stella Gagliardi[9], Lucia Corrado[11], Letizia Mazzini[12], Gianni Sorarù[13], Flavia Raggi[13], Gabriele Siciliano[14], Costanza Simoncini[14], Annalisa Lo Gerfo[14], Massimiliano Filosto[15], Maurizio Inghilleri[16,17], Alessandra Ferlini[18]

1. Department of Pathophysiology and Transplantation, “Dino Ferrari” Center, Università degli Studi di Milano, Milan, Italy
2. Department of Neurology and Laboratory of Neuroscience, Istituto Auxologico Italiano IRCCS, Milan, Italy
3. United Kingdom Dementia Research Institute, Maurice Wohl Clinical Neuroscience Institute, Department of Basic and Clinical Neuroscience, Institute of Psychiatry, Psychology, and Neuroscience, King's College London
4. Department of Medical Biotechnology and Translational Medicine, Università degli Studi di Milano, Milan, Italy
5. Unit of Medical Genetics and Neurogenetics, Fondazione IRCCS Istituto Neurologico Carlo Besta, Milan, Italy
6. 3rd Neurology Unit, Motor Neuron Diseases Center, Fondazione IRCCS Istituto Neurologico "Carlo Besta", Milan, Italy
7. Department of Medical Biotechnology and Translational Medicine, University of Milan, Milan, Italy
8. Neurology Unit, IRCCS Foundation Ca’ Granda Ospedale Maggiore Policlinico, Milan, Italy
9. Genomic and Post-Genomic Center, IRCCS Mondino Foundation, Pavia, Italy
10. Department of Brain and Behavioural Sciences, University of Pavia, Pavia, Italy
11. Department of Health Sciences, University of Eastern Piedmont, Novara, Italy
12. ALS Center Department of Neurology “Maggiore della carità” University Hospital Novara, Italy
13. Department of Neurosciences, University of Padova, Padova, Italy
14. Department of Clinical and Experimental Medicine, University of Pisa, Pisa, Italy
15. Department of Clinical and Experimental Sciences and NeMO-Brescia Clinical Center for Neuromuscular Diseases, University of Brescia, Italy
16. Department of Human Neurosciences, Rare Neuromuscular Diseases Centre, Sapienza University, Viale Dell'Università 30, 00185, Rome, Italy
17. IRCCS Neuromed, Pozzilli, Italy
18. Unit of Medical Genetics, Department of Medical Science, University of Ferrara, Ferrara, Italy

The members of Project MinE ALS Sequencing Consortium are:

Philip Van Damme[1,2], Philippe Corcia[3,4], Philippe Couratier[5,4], Patrick Vourc'h[6,7], Orla Hardiman[8,9], Russell McLaughlin[10], Marc Gotkine[11], Vivian Drory[12], Nicola Ticozzi[13,14], Vincenzo Silani[15,14], Jan H. van den Veldink[16], Leonard H. Berg[16], Mamede de Carvalho[17,18], Jesus S. Mora Pardina[19], Monica Povedano[20], Peter Andersen[21], Markus Weber[22], Ayşe Nazlı Başak[23], Ammar Al-Chalabi[24,25], Chris Shaw[26], Pamela J. Shaw[27], Karen E. Morrison[28], John E. Landers[29], Jonathan D. Glass[30]

1. KU Leuven - University of Leuven, Department of Neurosciences, Leuven, Belgium
2. VIB, Center for Brain & Disease Research, Laboratory of Neurobiology, Leuven, Belgium

3. Centre SLA, CHRU de Tours, Tours, France; UMR 1253, iBrain, Université de Tours, Inserm, Tours, France.
4. Federation des Centres SLA Tours and Limoges, LITORALS, Tours, France
5. Centre SLA, CHU Limoges, Limoges, France
6. Service de Biochimie et Biologie moléculaire, CHU de Tours, Tours, France
7. UMR 1253, Université de Tours, Inserm, 37044 Tours, France
8. Academic Unit of Neurology, Trinity College Dublin, Trinity Biomedical Sciences Institute, Dublin, Republic of Ireland
9. Department of Neurology, Beaumont Hospital, Dublin, Republic of Ireland
10. Complex Trait Genomics Laboratory, Smurfit Institute of Genetics, Trinity College Dublin, Dublin, Republic of Ireland
11. Department of Neurology, Hadassah Medical Organization and Faculty of Medicine, Hebrew University of Jerusalem, Israel
12. Department of Neurology Tel-Aviv Sourasky Medical Centre, Israel
13. Department of Neurology and Laboratory of Neuroscience, IRCCS Istituto Auxologico Italiano, Milano, Italy
14. Department of Pathophysiology and Transplantation, 'Dino Ferrari' Center, Università degli Studi di Milano, Milano, Italy
15. Department of Neurology and Laboratory of Neuroscience, Istituto Auxologico Italiano IRCCS, Milano, Italy
16. Department of Neurology, UMC Utrecht Brain Center, University Medical Center Utrecht, Utrecht University, Utrecht, The Netherlands
17. Instituto de Fisiologia, Instituto de Medicina Molecular,Faculdade de Medicina, Universidade de Lisboa, Lisbon, Portugal
18. Department of Neurosciences, Hospital de Santa Maria-CHLN, Lisbon, Portugal
19. ALS Unit, Hospital San Rafael, Madrid, Spain
20. la Unitat Funcional de Motoneurona, Cap de Secció de Neurofisiologia, Servei de Neurologia, Hospital Universitario de Bellvitge-IDIBELL
21. Department of Clinical Science, Neurosciences, Umeå University, Sweden
22. Neuromuscular Diseases Unit/ALS Clinic, Kantonsspital St. Gallen, 9007, St. Gallen, Switzerland
23. Neurodegeneration Research Laboratory (NDAL), Research Center for Translational Medicine (KUTTAM), Koç University School of Medicine, Istanbul, Turkey
24. Maurice Wohl Clinical Neuroscience Institute, Department of Basic and Clinical Neuroscience, Institute of Psychiatry, Psychology, and Neuroscience, King's College London
25. Department of Clinical Neuroscience, King's College Hospital, London SE5 9RS, UK
26. United Kingdom Dementia Research Institute, Maurice Wohl Clinical Neuroscience Institute, Department of Basic and Clinical Neuroscience, Institute of Psychiatry, Psychology, and Neuroscience, King's College London
27. Sheffield Institute for Translational Neuroscience (SITraN), University of Sheffield, Sheffield, UK
28. School of Medicine, Dentistry and Biomedical Sciences, Queen's University Belfast, UK
29. Department of Neurology, University of Massachusetts Medical School, Worcester, MA, USA
30. Department of Neurology, Emory University School of Medicine, Atlanta, GA, USA

The members of the American Genome Center are:

Adelani Adeleye[1,2], Camille Alba[1,2], Dagmar Bacikova[1,2], Clifton L. Dalgard[1,3], Daniel N. Hupalo[1,2], Elisa McGrath Martinez[1,2], Anthony R. Soltis[1,2], Gauthaman Sukumar[1,2], Coralie Viollet[1,2], Matthew D. Wilkerson[1,2,3]

1. The American Genome Center, Collaborative Health Initiative Research Program, Uniformed Services University of the Health Sciences, Bethesda, MD 20814, USA

2. Henry M. Jackson Foundation for the Advancement of Military Medicine, Bethesda, MD, 20817, USA
3. Department of Anatomy, Physiology & Genetics, Uniformed Services University of the Health Sciences, Bethesda, MD 20814, USA

Acknowledgments

The authors also acknowledge the ALS Association and the Muscular Dystrophy Association, which support DNA collection. This work was supported in part by the Italian Ministry of Health (Ministero dellaDella Salute, Ricerca Sanitaria Finalizzata, grant RF-2016-02362405, RF-2013-02355764, and GR-2016-02364373); the European Commission's Health Seventh Framework Programme (FP7/2007-2013 under grant agreement 259867); the Italian Ministry of Education, University and Research (Progetti di Ricerca di Rilevante Interesse Nazionale, PRIN, grant 2017SNW5 MB); the Fondazione Italiana di Ricerca per la SLA (AriSLA, Grant AZYGOS 2.0); the Joint Programme - Neurodegenerative Disease Research (Brain-Mend project) granted by Italian Ministry of Education, University and Research, and the Canadian Consortium on Neurodegeneration in Aging (CCNA). This study was performed under the Department of Excellence grant of the Italian Ministry of Education, University and Research to the 'Rita Levi Montalcini' Department of Neuroscience, University of Torino, Italy.

AAK is funded by the ALS Association Milton Safenowitz Research Fellowship (grant number 22-PDF-609.DOI:10.52546/pc.gr.150909.), The Motor Neurone Disease Association (MNDA) Fellowship (Al Khleifat/Oct21/975-799), The Darby Rimmer Foundation, and The NIHR Maudsley Biomedical Research Centre.

This study used DNA samples and clinical data from the Target ALS Human Postmortem Tissue Core, the NINDS Repository at Coriell, the North East ALS (NEALS) Consortium Biorepository, the New York Brain Bank-The Taub Institute, Columbia University, Department of Veterans Affairs Biorepository Brain Bank (grant #BX002466; CBB), the Baltimore Longitudinal Study of Aging (BLSA), and the Johns Hopkins University Alzheimer's Disease Research Center, the NICHD Brain, and Tissue Bank for Developmental Disorders at the University of Maryland.

The InCHIANTI study baseline (1998–2000) was supported as a "targeted project" (ICS110.1/RF97.71) by the Italian Ministry of Health and, in part, by the United States National Institute on Aging (contracts 263 MD 9164 and 263 MD 821336). The InCHIANTI follow-up 1 (2001–2003) was funded by the United States National Institute on Aging (contracts N.1-AG-1-1 and N.1-AG-1- 2111), and the InCHIANTI follow-ups 2 and 3 studies (2004–2010) were financed by the United States National Institute on Aging (contract N01-AG- 5-0002).

The data collection and genotyping of the *C9orf72* replication cohort at the King's College London was supported by MOTOR NEURONE DISEASE Association (MNDA) UK (grant code: Shaw/Apr18/864-791).

Project MinE data collection and genotyping of the *C9orf72* replication cohort was supported by funding provided by the Dutch Research Council (NWO) [VENI scheme grant 09150161810018] and Prinses Beatrix Spierfonds (neuromuscular fellowship grant W.F19-03 and by the Project number W.OR20-08 (The "Repeatome" as a basis for new treatments of ALS) of the Prinses Beatrix Spierfonds. and by the Dutch Research Council (NWO) [VIDI grant 91719350]. Project MinE is part of an EU Joint Programme - Neurodegenerative Disease Research (JPND) project. The project is supported through the following funding organizations under the aegis of JPND - www.jpnd.eu (the United Kingdom, Medical Research Council (MR/L501529/1; MR/R024804/1) and Economic and Social Research Council (ES/L008238/1)) and through the Motor Neurone Disease Association. This study represents independent research part-funded by the National Institute for Health Research (NIHR) Biomedical Research Centre at South

London and Maudsley NHS Foundation Trust and King's College London. This is, in part, an EU Joint Programme - Neurodegenerative Disease Research (JPND) project. The project is supported through the following funding organizations under the aegis of JPND - www.jpnd.eu (the United Kingdom, Medical Research Council (MR/L501529/1; MR/R024804/1) and Economic and Social Research Council (ES/L008238/1)) and through the Motor Neurone Disease Association. This study represents independent research part-funded by the National Institute for Health Research (NIHR) Biomedical Research Centre at South London and Maudsley NHS Foundation Trust and King's College London. In part, samples used in this research were obtained from the UK National DNA Bank for MND Research, funded by the MND Association and the Wellcome Trust. This project is partially supported by the AGING Project of the Department of Excellence at the Department of Translational Medicine (DIMET), Università del Piemonte Orientale, Novara, Italy. ALS Association Milton Safenowitz Research Fellowship funds AAK (grant number 22-PDF-609. DOI: 10.52546/pc.gr.150909.), The Motor Neurone Disease Association (MNDA) Fellowship (Al Khleifat/Oct21/975-799) and The NIHR Maudsley Biomedical Research Centre.

The dataset(s) used for the analyses described in this manuscript were obtained from the Age-Related Eye Disease Study (AREDS) Database found at https://www.nei.nih.gov/research/clinical-trials/age-related-eye-disease-study-aredsthrough dbGaP accession number phs000001.v3.p1. Funding support for AREDS was provided by the National Eye Institute (N01-EY-0-2127). We thank the AREDS participants and the AREDS Research Group for their valuable contribution to this research.

The Framingham Heart Study is conducted and supported by the National Heart, Lung, and Blood Institute (NHLBI) in collaboration with Boston University (Contract No. N01-HC-25195 and HHSN268201500001I). This manuscript was not prepared in collaboration with the Framingham Heart Study investigators. It did not necessarily reflect the opinions or views of the Framingham Heart Study, Boston University, or NHLBI. Funding to support the Omni cohort recruitment, retention, and examination was provided by NHLBI Contract N01-HC-25195 and HHSN268201500001I, and NHLBI grants R01-HL070100, R01-HL076784, R01-HL-49869, and U01-HL-053941.

Research support to collect data and develop an application to support this project was provided by 3P50CA093459, 5P50CA097007, 5R01ES011740, and 5R01CA133996.

The WHI program is funded by the National Heart, Lung, and Blood Institute, National Institutes of Health, and US Department of Health and Human Services through contracts HHSN268201600018C, HHSN268201600001C, HHSN268201600002C, HHSN268201600003C, and HHSN268201600004C. This manuscript was not prepared in collaboration with investigators of the WHI, has not been reviewed and approved by the Women's Health Initiative (WHI), and does not necessarily reflect the opinions of the WHI investigators or the NHLBI. Funding support for WHI GARNET was provided through the NHGRI Genomics and Randomized Trials Network (GARNET) (Grant Number U01 HG005152). The GARNET Coordinating Center assisted with phenotype harmonization, genotype cleaning, and general study coordination (U01 HG005157). The National Center assisted with data cleaning for Biotechnology Information. Funding support for genotyping, performed at the Broad Institute of MIT and Harvard, was provided by the NIH Genes, Environment and Health Initiative [GEI] (U01 HG004424). The datasets used for the analyses described in this manuscript were obtained from dbGaP at http://www.ncbi.nlm.nih.gov/sites/entrez?db=gap through dbGaP accession phs000001, phs000007, phs000187, phs000196, phs000200, phs000315, phs000675, phs000248, phs000292, phs000304, phs000368, phs000372, phs000394, phs000397, phs000404, phs000421, phs000428, phs000454, phs000615, phs000801, and phs000869.

The authors acknowledge the contribution of data from Hepatitis C Pathogenesis and the Human Genome supported by 1X01HG005271-01 and R01DA013324 and accessed through dbGAP to the analysis presented in this publication.

Funding support for the Genes and Blood Clotting Study was provided through the NIH/NHLBI (R37 HL039693). The Genes and Blood Clotting Study is one of the Phase 3 studies in the Gene Environment

Association Studies (GENEVA) under GEI. The GENEVA Coordinating Center assisted with genotype cleaning (U01 HG004446). Funding support for DNA extraction and genotyping, performed at the Broad Institute, was provided by NIH/NHLBI (R37 HL039693). The Howard Hughes Medical Institute provided additional support.

The dataset(s) used for the analyses described in this manuscript were obtained from the Genotype and Phenotype (dbGaP) database found at http://www.ncbi.nlm.nih.gov/gap through dbGaP accession number phs000368. Samples and associated phenotype data for the Genome-Wide Association Scan [GWAS] of Polycystic Ovary Syndrome Phenotypes were provided by Andrea Dunaif, MD.

The authors acknowledge the contribution of data from Genetic Architecture of Smoking and Smoking Cessation accessed through dbGaP. Funding support for genotyping, performed at the Center for Inherited Disease Research (CIDR), was provided by 1 X01 HG005274-01. CIDR is fully funded through a federal contract from the National Institutes of Health to The Johns Hopkins University, contract number HHSN268200782096C. The Gene Environment Association Studies (GENEVA) Coordinating Center (U01 HG004446) assisted with genotype cleaning and general study coordination. Funding support for collecting datasets and samples was provided by the Collaborative Genetic Study of Nicotine Dependence (COGEND; P01 CA089392) and the University of Wisconsin Transdisciplinary Tobacco Use Research Center (P50 DA019706, P50 CA084724).

The dataset(s) used for the analyses described in this manuscript were obtained from the Genetics of Fuchs' Endothelial Corneal Dystrophy (FECD) Study through dbGaP accession number phs000421. The grants that have funded the enrollment of the cases and controls to be used in this GWAS are R01EY016514 (DUEC, PI: Gordon Klintworth), R01EY016482 (CWRU, PI: Sudha Iyengar), and 1X01HG006619-01 (PI: Sudha Iyengar, Natalie Afshari). We would like to thank the FECD participants and the FECD Research Group for their valuable contribution to this research.

The authors acknowledge the contribution of data from the CIDR-NIDA Study of HIV Host Genetics accessed through dbGAP. Funding support for genotyping, performed at the Center for Inherited Disease Research (CIDR), was provided by 1 X01 HG005275-01A1. CIDR is fully funded through a federal contract from the National Institutes of Health to The Johns Hopkins University, contract number HHSN268200782096C. Funding support for the collection of datasets and samples was provided by NIDA grants R01DA026141(Johnson), R01DA004212 (Watters), U01DA006908 (Watters), R01DA009532 (Bluthenthal), the San Francisco Department of Public Health, SAMHSA, and HRSA.

The Genome-Wide Association Study (GWAS) of Non-Hodgkin Lymphoma (NHL) project was supported by the intramural program of the Division of Cancer Epidemiology and Genetics (DCEG), National Cancer Institute (NCI), National Institutes of Health (NIH). The datasets have been accessed through the NIH database for Genotypes and Phenotypes (dbGaP) under accession # phs000801. A complete list of acknowledgments can be found in the supplementary note (Berndt SI et al., Nature Genet., 201354).

This study made use of data generated by investigators in the BEACON consortium through a grant funded by the US National Institutes of Health (NIH) (RO1CA136725) to Thomas L. Vaughan and David C. Whiteman (multiple PIs). In support of this work, TLV was supported by NIH grant KO5CA124911 and DCW by a Future Fellowship grant FT0990987 from the Australia Research Council. Additional collaborators, sources of support, and origin of the data and biospecimens are listed in the following publication: Levine DM, Ek WE, Zhang R, Liu X, Onstad L, Sather C, et al. A genome-wide association study identifies new susceptibility loci for esophageal adenocarcinoma and Barrett's esophagus. Nat Genet. 2013 Dec;45(12):1487–93.

The American Genome Center is partly supported by an NHLBI grant: IAA-A-HL-007.001. The opinions and assertions expressed herein are those of the author(s). They do not necessarily reflect the official

policy or position of the Uniformed Services University, the Department of Defense, the Henry M . Jackson Foundation for the Advancement of Military Medicine, Inc., or the U.S. Government.

This project is partially supported by the AGING Project for the Department of Excellence at the Department of Translational Medicine (DIMET), Università del Piemonte Orientale, Novara, Italy.

RJBD is supported by the following: (1) NIHR Biomedical Research Centre at South London and Maudsley NHS Foundation Trust and King's College London, London, UK; (2) Health Data Research UK, which is funded by the UK Medical Research Council, Engineering and Physical Sciences Research Council, Economic and Social Research Council, Department of Health and Social Care (England), Chief Scientist Office of the Scottish Government Health and Social Care Directorates, Health and Social Care Research and Development Division (Welsh Government), Public Health Agency (Northern Ireland), British Heart Foundation and Wellcome Trust; (3) The BigData@Heart Consortium, funded by the Innovative Medicines Initiative-2 Joint Undertaking under grant agreement No. 116074. This Joint Undertaking receives support from the European Union's Horizon 2020 research and innovation program and EFPIA; it is chaired by DE Grobbee and SD Anker, partnering with 20 academic and industry partners and ESC; (4) the National Institute for Health Research University College London Hospitals Biomedical Research Centre; (5) the National Institute for Health Research (NIHR) Biomedical Research Centre at South London and Maudsley NHS Foundation Trust and King's College London; (6) the UK Research and Innovation London Medical Imaging & Artificial Intelligence Centre for Value-Based Healthcare; (7) the National Institute for Health Research (NIHR) Applied Research Collaboration South London (NIHR ARC South London) at King's College Hospital NHS Foundation Trust.

REFERENCES

S1. Brooks, B.R., Miller, R.G., Swash, M., Munsat, T.L., and World Federation of Neurology Research Group on Motor Neuron, D. (2000). El Escorial revisited: revised criteria for the diagnosis of amyotrophic lateral sclerosis. Amyotroph. Lateral Scler. Other Motor Neuron Disord. 1, 293-299. https://doi.org/10.1080/146608200300079536.

S2. Abrahams, S., Newton, J., Niven, E., Foley, J., and Bak, T.H. (2014). Screening for cognition and behaviour changes in ALS. Amyotroph. Lateral Scler. Frontotemporal Degener. 15, 9-14. https://doi.org/10.3109/21678421.2013.805784.

S3. Renton, A.E., Majounie, E., Waite, A., Simon-Sanchez, J., Rollinson, S., Gibbs, J.R., Schymick, J.C., Laaksovirta, H., van Swieten, J.C., Myllykangas, L., et al. (2011). A hexanucleotide repeat expansion in C9ORF72 is the cause of chromosome 9p21-linked ALS-FTD. Neuron 72, 257-268. 10.1016/j.neuron.2011.09.010.

S4. DeJesus-Hernandez, M., Mackenzie, I.R., Boeve, B.F., Boxer, A.L., Baker, M., Rutherford, N.J., Nicholson, A.M., Finch, N.A., Flynn, H., Adamson, J., et al. (2011). Expanded GGGGCC hexanucleotide repeat in noncoding region of C9ORF72 causes chromosome 9p-linked FTD and ALS. Neuron 72, 245-256. 10.1016/j.neuron.2011.09.011.

S5. Tunca, C., Seker, T., Akcimen, F., Coskun, C., Bayraktar, E., Palvadeau, R., Zor, S., Kocoglu, C., Kartal, E., Sen, N.E., et al. (2020). Revisiting the complex architecture of ALS in Turkey: Expanding genotypes, shared phenotypes, molecular networks, and a public variant database. Hum. Mutat. 41, e7-e45. https://doi.org/10.1002/humu.24055.

S6. Dolzhenko, E., van Vugt, J., Shaw, R.J., Bekritsky, M.A., van Blitterswijk, M., Narzisi, G., Ajay, S.S., Rajan, V., Lajoie, B.R., Johnson, N.H., et al. (2017). Detection of long repeat expansions from PCR-free whole-genome sequence data. Genome Res. 27, 1895-1903.https://doi.org/10.1101/gr.225672.117.

S7. Brooks, B.R. (1994). El Escorial World Federation of Neurology criteria for the diagnosis of amyotrophic lateral sclerosis. Subcommittee on Motor Neuron Diseases/Amyotrophic Lateral Sclerosis of the World Federation of Neurology Research Group on Neuromuscular Diseases and the El Escorial "Clinical limits of amyotrophic lateral sclerosis" workshop contributors. J. Neurol. Sci. 124 Suppl, 96-107. https:doi.org/10.1016/0022-510x(94)90191-0.

S8. van Rheenen, W., van der Spek, R.A.A., Bakker, M.K., van Vugt, J., Hop, P.J., Zwamborn, R.A.J., de Klein, N., Westra, H.J., Bakker, O.B., Deelen, P., et al. (2021). Common and rare variant association analyses in amyotrophic lateral sclerosis identify 15 risk loci with distinct genetic architectures and neuron-specific biology. Nat. Genet. *53*, 1636-1648. https://doi.org/10.1038/s41588-021-00973-1.

S9. Chang, C.C., Chow, C.C., Tellier, L.C., Vattikuti, S., Purcell, S.M., and Lee, J.J. (2015). Second-generation PLINK: rising to the challenge of larger and richer datasets. Gigascience 4, 7. 10.1186/s13742-015-0047-8.

S10. Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30, 2114-2120. https://doi.org/10.1093/bioinformatics/btu170.

S11. Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29, 15-21. https://doi.org/10.1093/bioinformatics/bts635.

S12. Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 15, 550. https://doi.org/10.1186/s13059-014-0550-8.

S13. Du, Z.W., Chen, H., Liu, H., Lu, J., Qian, K., Huang, C.L., Zhong, X., Fan, F., and Zhang, S.C. (2015). Generation and expansion of highly pure motor neuron progenitors from human pluripotent stem cells. Nat. Commun. 6, 6626. https://doi.org/10.1038/ncomms7626.

S14. van Rheenen, W., Shatunov, A., Dekker, A.M., McLaughlin, R.L., Diekstra, F.P., Pulit, S.L., van der Spek, R.A., Vosa, U., de Jong, S., Robinson, M.R., et al. (2016). Genome-wide association analyses identify new risk variants and the genetic architecture of amyotrophic lateral sclerosis. Nat. Genet. 48, 1043-1048. https://doi.org/10.1038/ng.3622.

S15. Wishart, D.S., Knox, C., Guo, A.C., Shrivastava, S., Hassanali, M., Stothard, P., Chang, Z., and Woolsey, J. (2006). DrugBank: a comprehensive resource for in silico drug discovery and exploration. Nucleic Acids Res. 34, D668-672. https://doi.org/10.1093/nar/gkj067.